

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

Etd

OM protein - protein search, using sw model

June 2011, 2005, 11.0.0.4.2, (search time 20 seconds) (without alignment)
3937.814 Million cell updates/sec

Title: AAC39336
Perfect score: 5677
Sequence: 1 MSEGETAGKPKMKNIVPQTC. 1065
Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters

283224

ALIGNMENTS

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%; Maximum Match 100%; Listing first 45 summaries

Database : PIR_73;*

1: pir1;*
2: pir2;*
3: pir3;*
4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Length	DB ID	Description
1	5677	100.0	T52054	cellulose synthase
2	3984	108.1	T05351	cellulose synthase
3	3811	67.1	T15179	probable cellulose synthase
4	3801	67.0	T0652	F84649
5	3628	63.9	T08494	cellulose synthase
6	3612.5	63.6	T152028	probable cellulose synthase
7	3590	63.2	T0882	H84604
8	3473.5	61.2	T10797	cellulose synthase
9	3317.5	58.4	T04870	cellulose synthase
10	2808	49.5	T10800	cellulose synthase
11	2324	40.9	T1812	hypothetical protein
12	2303.5	40.6	T1145	cellulose synthase
13	2289.5	40.3	T1111	hypothetical protein
14	2211.5	39.1	T05646	probable cellulose synthase
15	2185.5	38.5	T0361	probable cellulose synthase
16	1931	34.0	T202209	cellulose synthase
17	1902	19.2	T02561	probable cellulose synthase
18	1069	18.8	T712	cellulose synthase
19	1043	18.4	T02553	cellulose synthase
20	1018.5	17.9	T02560	cellulose synthase
21	953.5	16.8	T08918	hypothetical protein
22	948.5	16.7	T08920	hypothetical protein
23	939.5	16.5	T08919	hypothetical protein
24	934.5	16.5	E71417	hypothetical protein
25	849	15.0	B71417	hypothetical protein
26	441.5	7.8	A82275	cellulose synthase
27	305.5	5.4	T08591	hypothetical protein
28	301	5.3	T12093	TCACG motif bindin
29	283.5	3.2	T08592	TCACG motif bindin

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	100.0%	1065	T52054	cellulose synthase [validated]
2	100.0%	108.1	T05351	cellulose synthase
3	100.0%	67.1	T15179	cellulose synthase
4	100.0%	67.0	T0652	F84649
5	100.0%	63.9	T08494	cellulose synthase
6	100.0%	63.6	T152028	probable cellulose synthase
7	100.0%	63.2	T0882	H84604
8	100.0%	61.2	T10797	cellulose synthase
9	100.0%	58.4	T04870	cellulose synthase
10	100.0%	49.5	T10800	cellulose synthase
11	100.0%	40.9	T1812	hypothetical protein
12	100.0%	40.6	T1145	cellulose synthase
13	100.0%	40.3	T1111	hypothetical protein
14	100.0%	39.1	T05646	probable cellulose synthase
15	100.0%	38.5	T0361	probable cellulose synthase
16	100.0%	34.0	T202209	cellulose synthase
17	100.0%	19.2	T02561	probable cellulose synthase
18	100.0%	18.8	T712	cellulose synthase
19	100.0%	18.4	T02553	cellulose synthase
20	100.0%	17.9	T02560	cellulose synthase
21	100.0%	16.8	T08918	hypothetical protein
22	100.0%	16.7	T08920	hypothetical protein
23	100.0%	16.5	T08919	hypothetical protein
24	100.0%	16.5	E71417	hypothetical protein
25	100.0%	15.0	B71417	hypothetical protein
26	100.0%	7.8	A82275	cellulose synthase
27	100.0%	5.4	T08591	hypothetical protein
28	100.0%	5.3	T12093	TCACG motif bindin
29	100.0%	3.2	T08592	TCACG motif bindin

C;Keywords: glycosyltransferase; hexosyltransferase
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>
A;Cross-references: EMBL:AF027174; PIDN: AAC39336.1
A;Genetics:
A;Note: Ath-B
C;Function:
A;Description: EC 2.4.1.; cellulose synthase [validated, MUID:98111412]; involved in cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C;Accession: T52054
R;AxioLi, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Camilleri, C;Reference number: 213745; MUID:98111412; PMID:9445479
A;Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1065 <ARI>

QY 2 ESGETAGKPMENIVPOTQCISDNNGKTVDODRFACDTSFPVRCYERKGNO 61
 421 IDYLDKVQTSEVKDRRAKREVEEKFIRINALYSKALKCPEGWMDGPWPNNTGD 480
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 421 IDYLKDVKYTSEVKDRRAKREVEEKFIRINALYSKALKCPEGWMDGPWPNNTGD 480
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 481 HPOMIQFLGONGGLDAEGNEPLRVYVSRERKPGFOHKKAGAMNALVRSAVLNGPF 540
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 481 HPMIQVFLGONGGLDAEGNEPLRVYVSRERKPGFOHKKAGAMNALVRSAVLNGPF 540
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 541 IUNLDCDORYINNSKALRREAMCFIMDPNLGKOCYVOPRERGIDKNDRYAARNRTF 600
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 541 IUNLDCDORYINNSKALRREAMCFIMDPNLGKOCYVOPRERGIDKNDRYAARNRTF 600
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 601 NLRGDLG3IQQGPPVYGTGCVFNRTALYGYEPPIKVHKKPSLSSKLCCGSKKAKKES 660
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 601 NLRGDLG3IQQGPPVYGTGCVFNRTALYGYEPPIKVHKKPSLSSKLCCGSKKAKKES 660
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 661 DKKSGRHTDSTPVVFNIIDIEGVEGAGFDDEKALLMSQMLEKREGQSAVFASTLME 720
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 661 DKKSGRHTDSTPVVFNIIDIEGVEGAGFDDEKALLMSQMLEKREGQSAVFASTLME 720
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 721 NGGVPPSATPENILKEATHVSGCYEDKSDWMEIGMIGYGSYTEDILTGKMHARWSR 780
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 721 NGGVPPSATPENILKEATHVSGCYEDKSDWMEIGMIGYGSYTEDILTGKMHARWSR 780
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 781 YCMPLKPAFKGSAPINSLDRNQVLRWALGSVEILSFRHCPITWGYNGRLKFERAYVN 840
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 781 YCMPLKPAFKGSAPINSLDRNQVLRWALGSVEILSFRHCPITWGYNGRLKFERAYVN 840
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 841 TTYPITSPILLMCTLLAVCLFTNQIPIQNSIASIWLSLFLSIFATGLEMRSVG 900
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 841 TTYPITSPILLMCTLLAVCLFTNQIPIQNSIASIWLSLFLSIFATGLEMRSVG 900
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 901 GIDEMWRNEQFWWIGGYSAHALFAVFOGLTILKLAGIDTNFTVTKSASBDGDFAEYLFWK 960
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 901 GIDEMWRNEQFWWIGGYSAHALFAVFOGLTILKLAGIDTNFTVTKSASBDGDFAEYLFWK 960
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 961 TLLIIPPTPLTLNVLYGVAGYSAINGSYOSWGPLGKLFPAFWTVHLYPFLGLMR 1020
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 961 TLLIIPPTPLTLNVLYGVAGYSAINGSYOSWGPLGKLFPAFWTVHLYPFLGLMR 1020
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 1021 QHPTPYWWSVLLASIFSLIWWRIOPFTSRVTPDILECINC 1065
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 1021 QHPTPYWWSVLLASIFSLIWWRIOPFTSRVTPDILECINC 1065
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 RESULT 2
 T05351 cellulose synthase (EC 2.4.1.-) catalytic chain RSW1 - Arabidopsis thaliana
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 N; Alternate names: protein F8B4_110
 C; Species: Arabidopsis thaliana (mouse-ear cress)
 C; Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
 C; Accession: T05351
 R; Bevan, M.; Terryn, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.; De ewes, H.W.; Mayer, K.F.X.; Schueler, C. Sequence Database, February 1999
 A; Reference number: Z15409
 A; Accession: T05351
 A; Molecule type: DNA
 A; Residues: 1-1081 <REV>
 A; Cross-references: EMBL:Al034567
 A; Experimental source: cultivar Columbia; BAC clone F8B4
 C; Genetics:
 A; Gene: RSW1
 A; Map position: 4
 A; Introns: 27/3; 93/1; 150/3; 189/1; 252/2; 341/2; 456/3; 502/3; 544/3; 615/3; 704/1; 76
 C; Notes: F8B4_110
 Query Match Score: 3984; DB: 2; Length: 1081;
 Best Local Similarity: 69.7%; Pred. No. 4; ge-307;
 Matches: 741; Conservative: 128; Mismatches: 158; Indels: 36; Gaps: 13;

		RESULT 3	
T51579		cellulose synthase catalytic subunit (IRX3) - Arabidopsis thaliana	
N	Alternate names:	protein T10B6_80	
C	Species:	Arabidopsis thaliana (mouse-ear cress)	
C	Date:	18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000	
C	Accession:	T51579	
R	Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asanizu, E.; Kotani, H.; Tabata, S.; Mew		
A	Submitted to the Protein Sequence Database, August 2000		
A	Reference number:	225394	
A	Accession:	T51579	
A	Status:	preliminary	
A	Molecule type:	DNA	
A	Residues:	1-1026 <SAT>	
A	Cross-references:	EMBL:AU391142	
A	Experimental source:	cultivar Columbia; BAC clone T10B6	
C	Genetics:		
A	Map position:	5	
A	Introms:	25/3; 91/1; 184/1; 215/2; 304/2; 419/3; 507/3; 578/3; 650/1; 715/3; 833/3	
A	Note:	T10B6_80	
Query Match	67.1%	Score 3011; DB 2; Length 1026; Best Local Similarity 66.9%; Pred. No. 2.2e-293; Matches 721; Conservative 115; Mismatches 143; Indels 98; Gaps 17; Qy	
10	KPMKNIVPOTCQICSDNNVGTVDGDRFWACDIOSSPPVRCYEVERKDGQNQSCPOCKRY	69	
27	KPLKNLDSOFCETCGDQIGLTVEGLFACNECGFPACRCPYEVERRGTCNQPCCKRY	86	
Qy	70	KRLKGSPALPGDKDEGDGLAD--EGTVENPYOKEKIERMLQWHLTGKGEIME-QY	125
Db	87	KRLKGSPRVEGDEDDEDDIDIEEFNIEHEQDKHHSAAEMLYKGMSTGRGPEDDENGKF	146
Qy	126	DKEVSHNHPRLSRQDTSGEFASASPRPLSVSSTAG-----KRL_PYSDDNQ	175
Db	147	PPIVAGGH-----SGEF-----PVGGYNGEHHGLRKRVHYPPSSAG	184
Qy	176	SPNRRIIVDPVGGLGNVAKWERGVDGKMKQEKNTGPVSTQAERGVVIDASTDILADEL	235
Db	185	SEG-----GWRERMDDWKL-GHGNLGPPEP-----DDPEMELI-----	216
Qy	236	LNDEARQPLRSKRVISIPSSRINPVRNVIMLVLVCLFLHYRITNPVPNPAFALNLVSVICE	295
Db	217	--DEARQPLSKRPIASNSKINPVRNVIARLVLVLAFLYRILNPVHDLGLNLVSVICE	274
Qy	296	IWFALSWLQDFRKPWFNPRTYDRLALRYDREGEPQLAAYDIFSVTVDPEKPPVT	355
Db	275	IWFAVSWLQDFRKPWFNPRTYDRLALRYDREGEPQLAAYDIFSVTVDPEKPPVT	334
Qy	356	ANTVLSILAVDYPVYDGAAMLSFESLAETSEPARKWPFCKKYLEPRAEW	415
Db	335	SNTVLSILAMDYDPEVKCISCVYSDGASHMLTFESLSETAEFAKRWMPFCKKFESTPRAEW	394
Qy	416	YFAAKIDIKDKVQTSFKDRRANKREVEEFKRIRNALVSKALKCPEGWVMODGTPHPG	475
Db	395	YFTLKVDLQDKVHPTFVKERRAKREVEEFKRIRNALVSKALKCPEGWIMQDGTTPHPG	454
Qy	476	NNTGDDHPSGMTOVFLGONGSLDAGSNELRLVVYSSREKPGFOHKKACAMNALVRSVL	535
Db	455	NNTKDHPGSMIQVFLGONGSLDAGSNELRLVVYSSREKPGFOHKKACAMNALVRSVL	514
Qy	536	TNGPFILMLDCDHVYINNSKALREAMCFLMPNQPKQVYQFOPRFDDIDKNDRYANRNT	595
Db	515	TNAPFMNLDCDHVYINNSKALREAMCFLMPNQPKQVYQFOPRFDDIDKNDRYANRNT	574
Qy	596	VFFDINLRQLDGJQGPVYVGTGVFVNFRALYGEPPPIKVKKHCPSSLKLCG----GSR	650
Db	575	VFFDINMGQDGLQGPVYVGTGVFVNFRALYGEPPPIKVKKHCPSSLKLCG----GSR	630
Qy	651	KKNNAKAKEDDKKSGRHTDSTPVENFDLIEDIECSEG-----AFGDEKDALKLMSQMLEKRGQ	709
Db	631	RKNKK-----FSKNDMNGDAVALLGAEGSKREHJMSEMFEKTGQ	670
Qy	710	SAVIVVASTMENGVPPSATPENILKEATHVISGYEDKSDWMGEIGWTGYSVTEIDLNG	769
Db	671	SSIVVTSTLMEEGVPPSSPAVILKEATHVISGYEVTEWLGEVSTEDILT	730
Qy	770	FKMARGWHSIYCKPLPKAFKGSPAINISDRDNQVLRALGSYETEFLFRHCPWYGY	828
Db	731	FKMHCGRWRSIYCKPLPKAFKGSPAINISDRDNQVLRALGSYEVTEFLFRHCPWYGY	790
Qy	829	RLKLERAYVNTTYPITSIPISLUMCYTLLAVGIFTNOFIQPSINASIWEFLSIF	888
Db	791	KLKWLERAYVNTTYPITSIPISLUMCYTLLAVGIFTNOFIQPSINASIWEFLSIF	850
Qy	889	ATGILEMRNSVGVGDEWWRNQEWVIGVSAHLFIAFVFOGILKVLLAGIDNTFTVSKASD	948
Db	851	VTGELERLNSGVSTEWNRNEQFNVIGGISAHLFIAVQCLKLLAGIDNTFTVSKATDD	910
Qy	949	DGDPAELYAFKWTLLIPPTLLVNLVLYGVWVAGSYAINGSKYOSWGLPFGKLFCAFNV	1008
Db	911	D-DGELTAFKWTLLIPPTVILININGVWAGISDAINGTQSGPLFGKLFASFNV	969
Qy	1009	HYFLKLGLMRQRTPTVWVMSVLLASIFSLWVDRPFSRVTGPDILEGGINC	1065
Db	970	HYFLKLGLMGROMRRTPTVWVMSVLLASIFSLWVDRPFLVLTGKPDTSKCGINC	1026
Qy	RESULT 4		
Qy	884649	Probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana	
C	Species:	Arabidopsis thaliana (mouse-ear cress)	
C	Date:	02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001	
C	Accession:	F84649	
R	Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.		
M	Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, M.; Neiman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999		
Nature	402,		
A	Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.		
A	Reference number:	AB4420; PMID:20083487; PMID:10617197	
A	Accession:	F84649	
A	Status:	preliminary	
A	Molecule type:	DNA	
A	Residues:	1-1065 <STO>	
A	Cross-references:	GB:AE002093; NID:94432865; PIDN:AAD20713.1; GSDB:GN00139	
A	Gene:	At2g25540	
A	Map position:	2	
Query Match	67.0%	Score 3001; DB 2; Length 1065; Best Local Similarity 66.2%; Pred. No. 1.5e-292; Matches 714; Conservative 132; Mismatches 156; Indels 76; Gaps 12; Qy	
10	KPMKNIVPOTCQICSDNNVGTVDGDRFWACDIOSSPPVRCYEVERKDGQNQSCPOCKRY	69	
Qy	23	KPLKNLDSOFCETCGDQIGLTVEGLFACNECGFPACRCPYEVERRGTCNQPCCKRY	82
Db	70	KRLKGSPALPGDKDEGDGLAD--EGTVENPYOKEKIERMLQWHLTGKGEIME-QY	129
Db	83	RHNGSPRVEVEDKDDNDIEN-EFDYQ-----	111
Qy	130	SHNHLPLRTSQRDTSGERSAAS--PERUSVSTTAGKRLPSSDVNOSPNR-----	179
Db	112	-GNKIKARLPHAE--EISSSRHEEUSPVSLTHGH--PVSGEI-PTPDRNATLSPCI	163
Qy	180	-----RIVD-----VSGIVNVAKWERGVDGKMKQEKNTGPVSTQAER	218
Db	164	DQQLPGIYQLLLEPVLRILPSKDLINSYGLVNWDKKR1QGWKLQDKNMHMGKYHGK	223
Qy	219	GGVIDASPDILADEALUNDEASQPLSKVYSPSRNPYRMVIMRLVILCFLFHRYT	278
Db	224	GGE---FEGTGSNGDLDQMDYDARLPSRVMHPPSAPMTYRPTVILRLLGFLVFLHRT	281
Qy	279	NPVNPNAFLWLVSVICEWFWSWLDQFPKWPVNPRTYDRLALRYDREGESQAV	338
Db	282	HPPVFDAYALWLTSVICETWAFWSWLDQFPKWPVNPRTYDRLALRYDREGESQAV	341

Db 909 TGILEMQGGVGDMMWRNEOFWVGGASSHLFALFOGLKLKVLAGVNTNFTVSKAD-D 967
 Qy 950 GDFAEILYLFKMTTLLIPPTTLLIVNLUVGWAGVSYAINSGQSWGLFLRGKJEFKAWVHV 1009
 Db 968 GAFSELYLFKMTTLLIPPTTLLINIGVGSVDASWGPFLERGFLFAWVHV
 Qy 1010 LYPFLKGMLMGQRNRTPTIVVWSVILLASIFSLWVRIDPFTSRVTGPDILRCGINC 1065
 Db 1028 LYPFLKGMLGKDMPKTTIVVWSILLASILTLWVWVNPVFAK-GGPVLEICGLNC 1082
 RESULT 6
 TS2028 cellulose synthase [imported] - *Arabidopsis thaliana* (fragment)
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C;Accession: TS2028
 R;Joshi, C submitted to the EMBL Data Library, May 1998
 A;Reference number: Z25890
 A;Accession: TS2028
 A;Status: preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: mRNA
 A;Residues: 1-1081 <JDS>
 A;Cross-references: EMBL:AF062485; PIDN: AAC29067.1
 Query Match 63.6%; Score 3612.5; DB 2; Length 1081;
 Best Local Similarity 63.5%; Pred. No. 1.3e-277;
 Matches 685; Conservative 165; Mismatches 183; Indels 45; Gaps 21;
 Qy 10 KPMKNITVPTQTCOICSDNVGKTVGDDRFVACDICSFPVCPCYEYERDGNOSCPQKTY 69
 Db 26 RSVQELSGQTCQICRTELTWDPFVACNECAFVCPYERYERREGNQACPQCKTRF 85
 Qy 70 KRLKGSPPAIPEGKDEGLADEGTVEFNYPQK---EKISERMGLWHLTRGKGEEMKEPQY 125
 86 KRLKGSPPAIPEGKDEGLADEGTVEFNYPQK---EKISERMGLWHLTRGKGEEMKEPQY 138
 Qy 126 DKEVS--HHHPPLTSRQDTSGEFSASPERLVSSTIAG-GKRL-PYS-SDVNQSPNRR 180
 Db 139 DLDASPQGSQLPLT-YGDEDEVIS-SDRHALIVPSLGGIGNRVHPSLSDPTVAHR 196
 181 IVDP---VGLGNPAKERVGDKWMKQEKNTGPVSTQAAERGGVVDIDASTDILADEA 234
 Qy 197 LMVPOKDLAVYGGSYAWKDMEEWKKRQNE---KLQVRHREGBPFDFEGD--ADFP 249
 235 LLNDEARQLPSRKVKVSPSRINPYRAMILRLVILFLHYRITNVPNAFLWLVVIC 294
 250 MM-DEGRQPLSKMKIPKSKINPYRMLVLRLVILGLFHFYRILHVKDAYALWLWVIC 308
 Qy 295 EIWFASWILDOEPKWPVNPNEYTYLRLALYDREGEPSQLAADFVYSTWDPLKPPV 354
 Db 309 EIWFASWILDOEPKWPVNPNEYTYLRLALYDREGEPSQLAADFVYSTWDPLKPPV 368
 Qy 355 TANTVLSILANDYPVOKVSCVFDDGAMLSFSLAETSEPARWKWPCKKYSIERAPE 414
 Db 369 TANTVLSILANDYPVOKVACTVSDGCAALMFEAUSETAAEARKWPCKKYSIERAPE 428
 Qy 415 WFAAKIDYLKDVKQTTSFKDRRAMKREYEFKIRINALVAKKCPPEGMQMDCTPWP 474
 429 WYCFKMDYLKRNKVKHFAFRERRAMKRDYEEFKVKNALVATAQKVPEDGTMQDGTPWP 488
 Qy 475 GNNTGHPGPMTOVEUQNGGDAEGHPELPRUVYVSEKRCFOHKKAGANALYRVSAY 534
 Db 489 GNSVRDQHGPMLQVELGSDGVRODNEHPLRPUVYVSEKRCFOHKKAGANSLTRVSGV 548
 Qy 535 LTNGPFIILNCDHYINNSKAIRREAMCFLMPNLGKQCVYQFQRFIDKNDRYARN 594
 Db 549 LSNAPYILNVDCHYINNSKAIRREAMCMMQPSQSKRICKVQFQRFIDRHDYRSNR 608
 Qy 595 TVFDDINLRGLDGIQGPVYVQTCGVNTRALYGYERPIKVKHKKPSL----LSKUCGS 649

Db 609 VVFDINNKGLDGLQGPIYVGTCVFRQQALYGFDAKKKKGPKRTCNCKWPKWCLLCFGS 668
 Qy 650 RKKNSKAKK-ESDKKSSGRHDTSTPVFNLDIEVEGAGFDDERKALLSMSLSKRCF 708
 Db 669 R-KNRKAKTVAADKKKR-N-EASKOJHALENIEGRGHKVLYNVESTEAMOMKQKKG 725
 Qy 709 QSAVFAVASTLMENGVPSATPENLKEAIHWISCGYEDKSDWMEGTGWYGSVTDILT 768
 Db 726 QSPFVVASRLNGGMARNASPACLIKEIQVISGYEDKTEWKGKEIGWYGSVTDILT 785
 Qy 769 GFRMHARGWRSYCMPLKUPAKGSAPINSDRNQVRWALGGSVELFSRCPWNYGY 828
 Db 786 GSKMWHGWRHRYCTPKLAFFKGSAPINSDRNLHOTLWALGGSVELFLSRCPWNYGY 845
 Qy 829 RIKFLERFAYVMTTIVTISPLIMCFLAVCLTNQFIPQINSIASFLSLSIF 888
 Db 846 GLKWLRLSUYINSVSVVPPWTSLPLIVVCSLPATCILACKFIVPEISNYASILEMFLASSIA 905
 Qy 889 ATGILEMRWASGYGIDEWARNEOFWVLLGVSALHAFVAFQGILKVALGIDTNTVTSKASDE 948
 Db 906 ITGLEMOWGKVGIDDMWRNEQFWVIGVGSALFLFOGLKVALGQDNTVTSKASDE 964
 Qy 949 DGDFAEILYLFKMTTLLIPPTTLLIVNLUVGWAGVSYAINSGQSWGLFLRGKJEFKAWVHV 1008
 Db 965 DGEFSDLYLFKNTSLLIPPTTLLIVNLUVGWAGVSYAINSGQSWGLFLRGKJEFKAWVHV 1024
 Qy 1009 LYPFLKGMLMGQRNRTPTIVVWSVILLASIFSLWVRIDPFTSRVYGPDLI-CGJNC 1065
 Db 1025 HLYPFLKGMLGKDMPKTTIVVWSILLASILTLWVNPVFAK-GGPFILEICGLDC 1080

RESULT 7
 H84604 probable cellulose synthase catalytic subunit [imported] - *Arabidopsis thaliana* (mouse-ear cress)
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C;Accession: H84604
 R;Lin, X.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: H84604
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1088 <JDS>
 A;Cross-references: GB:AE002093; NID:9441721; PIDN:AAD20396.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: At:g21770
 A;Map position: 2
 Query Match 63.2%; Score 3590; DB 2; Length 1088;
 Best Local Similarity 62.5%; Pred. No. 8.3e-276;
 Matches 671; Conservative 163; Mismatches 191; Indels 48; Gaps 16;

Db 18 ORCOCISDNVGTWVGDIDRFVACDICSFPVCPCYEYERDGNOSCPQKTRYKRGSPA 77
 Db 37 QPKCKICRDEIELTDNGEFPFIACNECAPTCRCPYERYERREGNQACPQCGTTRYKRGSPR 96
 Qy 78 IPGDKDEGLADEGTVEFNYPQEKESERMGLW-HLTRGKGEEMGERQDVKEVSH---- 131
 Db 97 VEGDEDDI-DDLEHEFHYGMDPHEVTEALYYMNRGTD----EVSYLAYS 147
 Qy 132 -NHLPLTSRQDTSGEFSASPERLVSSTIAGIR--IPIYSDVNQSPNRRIVDP-- 184
 Db 148 PGSEPVPLTYCDEDSDMS--DRHALIVPPGGLGNVRHVPFTSFASHTRPMVQKD 205
 Qy 185 --VGIGNVAKKERVDPCKWKMKEKNTGPVSTQAAERGGVVDIDASTDILADEL-LNDEA 240
 Db 206 LTVYGYGSVAKDRMEWKKQOIKEQLQVVKNERVNDGDDGF--IVDELDOPGPGLPMDEG 263
 Qy 241 RQPLSRKVSPSRINPYRMVIMLRLVILFLHYRITNVPNAFLWLVVICIWAL 300

A;Gene: celA1
 C;Function:
 A;Description: involved in the synthesis of cellulose
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 61.2%; Score 3473.5; DB 2; Length 974;
 Best Local Similarity 62.1%; Pred. No. 1.2e-266;
 Matches 660; Conservative 129; Mismatches 176; Index 97; Gaps 15;

QY 12 MKNIVPQQCICSDNGKTVGDGRFVACDICSPPCVCYERKDNOSCPOCKTNYKR 71
 Db 1:: | - | : || | : || | : || | : || | : || | : || | : || | : || | : || |
 2 MESSCP-VCHTCBEBHVGAVNGGPFVACHECNPCICKSCFEDLKREGKACIRC----- 54
 QY 444 MDYIKHKUDPAFMERRAKRDKDEEFKVKINALVSKALKPERCQVMWMDGTPWPGNNNGD 480
 Db 481 HPGMLOQEVFGONGSGLDAEKGELRLRVVYSSREKPGFQHKKAGAMNALVRVSAYLTNGP 503
 QY 504 HPGMLOQEVFGHSGVCDMONGNELPLVYSSREKPGFQHKKAGAMNALVRVSAYLTNGP 540
 Db 541 IILUDCDHYINNSKALREAMCFLMDPNLGKQVYQVFQRFDGIDKNDRYANRNTVFDI 600
 QY 564 LLNVDCDHYINNSKAIRAFMCFLMDPQSGKICVQFQRFDGIDHRHRSNRNWFID 623
 QY 601 NLRLDGGTQGPVYVGTCVFNRTALYGEPPIKVHKHP----SLISK--LGGSRK 652
 1:: || | : || | : || | : || | : || | : || | : || | : || | : || | : || | : || |
 624 NMKGLDGIGQPIVGTGCVFRQALYGFADP--KKQGPGRCTNCWPWCCCLCGMRK 680
 QY 653 NSRAKKESDKKSGSRHTSTVPUENLDIEEGVAGAGTEDEKALLMSOMSLERKFQGSAY 712
 Db 681 KTGVKVDNOKRK---KESKQTHALEHEEGIQTNAENSE---TADQKLKEKFGOSPV 735
 QY 713 FVASTLMENGGVPSATPENNLKEAIRVTSCGEDKSDWGMETGWIGSVTEILTGRKM 772
 Db 736 LVASTLLINGGVPSVNVPNRLRESIOVSCGYEEKTEWGKEIGWIGSVTYEDLTGRKM 795
 QY 773 HARQWSRIVYCOMPKLAFKGSGAPINLSDRINQNVRWALGSVEILFSRHOPPIWYCYNGRIK 832
 Db 795 HCHQWRSVCMPPRAFKGSAPINLSDRHQVJRWALGSVEILFSRHOPPIWYCYGGGIKW 855
 QY 833 LERFAYVNTTIPPTSIPLMCTTLLAVCLFLTNQFTIPOISNIATSWLISIFANGI 892
 Db 855 LERFSYINSVVIPTSLPLLIVYSLPSLPAICLTLGKFIPEISNAGILEMENSIAVTG 915
 QY 893 LEMRWSGVYIDEMWRNEOFFWWIGGSVAHFAVOFGIQLKVLAGIDTNFTWTSKASDGF 952
 Db 916 LEMWQGKICIDDMWRNEOEWVWIGGVSHLFALFOGLKLWVLAGYSTNTFTWTSKAD-DGF 974
 QY 953 AELYFKWTTLIPPTTLLIVLNQWVAGSYVAINSGQSQWGLFLGKLEFAFWIVHLYP 1012
 Db 975 SELYFKWTTLIPPTTLLIVLNQWVAGSYVAINSGQSQWGLFLGKLEFAFWIVHLYP 1065
 QY 1013 FLKGLMGRONRPTIIVVWVSLASIFSLWLWTRDPFSTSVGDILECGINC 1065
 Db 1035 FLKGLLGKDRPVIILWVSLASILTLLWVRFNPVSK-DGPVLECGLDC 1086

RESULT 8

T10797 cellulose synthase (EC 2.4.1.-) catalytic chain celA1 - upland cotton
 C;Species: *Gossypium hirsutum* (upland cotton)
 C;Accession: T10797
 R;Pearl, J.R.; Kawagoe, Y.; Schreckengost, W.E.; Delmer, D.P.; Stalker, D.M.
 Proc. Natl. Acad. Sci. U.S.A. 93, 12637-12642, 1996
 A;Title: Higher plants contain homologs of the bacterial celA genes encoding the catalyst
 A;Reference number: 217152; MUIR:97057296; PMID:8901635
 A;Accession: T10797
 A;Status: preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: mRNA
 A;Residues: 1-974 <PEA>
 A;Cross-references: EMBL:U58283; NID:91706955; PIDN:AAB37766.1; PID:91706956
 A;Experimental source: strain Acala SJ-2; fiber
 C;Genetics:

QY 966 PPTTLIVNLYGWAGWSYALNSGYQSWGPFLGKLFPAFWVIVHLYPLFKGLMGRQRRTP 1025
 |||||||:|||||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 873 PPTTLIVNMYGVAGFSDALNKGYEAWGPLGKVFFSWTILHYPLFKGLMGRQRRTP 932

QY 1026 TIVVMSVLLASIFSLIWNRIDPFTSRVTPGDL--CGIN 1065
 |||||||:|||||:|||||:|||:
 Db 933 TIVVMSVLLASIFSLIWNRIDPFTSRVTPGDL--CGIN 974

RESULT 9

T04870 cellulose synthase (EC 2.4.1.-) catalytic chain F28A21-190 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
 C;Accession: T04870
 R;Bevan, M.; Mueller, M.W.; Muellein, A.; Feiber, R.; Bancroft, I.; Mewes, H.W.; Mayer, A.; Reference number: Z15387
 A;Accession: T04870
 A;Molecule type: DNA
 A;Residues: 1-938 <BREV>
 A;Cross-references: EMBL:AL035526
 A;Experimental source: cultivar Columbia; BAC clone F28A21
 C;Genetics:

A;MAP position: 4
 A;Introns: 60/1; 67/3; 94/1; 149/2; 240/2; 355/3; 401/3; 443/3; 514/3; 662/3; 780/3; 791
 A;Note: F28A21-190
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 58.4%; Score 3317.5; DB 2; Length 958;
 Best Local Similarity 60.6%; Pred. No. 2.8e-254; DB 2;
 Matches 639; Conservative 124; Mismatches 173; Indels 119; Gaps 16;

QY 20 COICSDWKGKTYGDGFREACIDCSFPVPCYEVERKGNGNOSCPCQKTRKRLKGSPAI 79
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 9 CMCGEIGVNGSERFVACHCSEPCPKACLEYREGRCLRC------GPAI 57

QY 80 GDKDEGLSLADECTVEFNYPQEKISERMLGWLHTRGKGEEEMGEPOYDKEVSHNLPRLT 139
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 58 -----IDENV-FDDVETTS 71

QY 140 RODTSGEFSASPERISVSSTAGKRLPYSSDVNOSPNRRLVDPVGLIGNYAWKERVDW 199
 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 72 K-TQSIDSGI-HARHISTVSI-----DSELNDE-----YGNPWTKNRVEW 111

QY 200 KMKOKEW-----GPVSTQMASERGGYDIDASTDILADEALINDE--ARQPLSKRVS 252
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 112 KDKDKKKKKKKKKDKPKTAKEHQAOQ---PTQHMEDTPNTESGATDVLWVPIR 167

QY 253 SRINPYRMVIMLRIVLICLFLHYRINPPNAFALWLVSICEIWALSWILDQFPKWP 312
 ::| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 168 TRKTSYRIVVIMRLITLALLENFYRITHPVDSAYGLWLTSCIEIWASWVLDQFPKWP 227

QY 313 VNRRETIDRLALRYDREGEPSOLAANDIFVSTVDPKLEPPLYNTAVSILAVDYPVDK 372
 :|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 228 INRETYIDRLSARFERPEEQSOLAANDFVSPDKEPLLTANTVSLIALDYFDKV 287

QY 373 SCYVFDDGAAMLSFSFESLAETSFARKWVPECKYSTERPAPWYEAKKIDYLKDVKOTSE 432
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 288 SCYVSDGGAMLSFSFESIYETADFAKKWVPECKYSTERPAPWYEAKKIDYLKDVKOTSE 437
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 433 VKDRRAMKREEFKIRINALVSKALKCPECGWMDGTPWPGNNQDHGPMIQVQLQN 492
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 348 VKERRAMKRDYBEFKIRMNALVAKAOKTPCPEBGTMQDGTSWPNNTFRDHPGMIOVFLGS 407

QY 493 GOLDAEGNELPVLVYSREKRGFQHHKAGAMNALVRSAYLTNGPFILNLDCDIN 552
 :|||||:|||||:|||||:|||||:|||||:|||||:
 Db 408 GARDISGENELPVLVYSREKRGFQHHKAGAMNALVRSAYLTNGPFILNLDCDIN 552
 :|||||:|||||:|||||:

QY 553 SKALREAMCFMLDPNPKGQVYCVOFRDFGTDKNDRYANRNTVERDINLGLDGQGPV 612
 :|||||:|||||:
 Db 468 SKAVREAMCFMLDPVQGDVCFVOPFRDFGTDKSDRYANRNTVERDINLGLDGQGPV 527

QY 613 YWGTCVNRITALGYEPPIKVKKKPSLLSKLGGSRKKNSKAKKESDKKKSGRHSDT 672
 :|||||:|||||:
 Db 528 YUGTGTVERROAQGYGSPSKPRLPQSSSSCCCLTKKK--OPDSESYKDAERL 585

QY 673 VPVNLDDEEGVAGEGDD-EKALLMSQSLERKFGOSAVFASLTMENGVPPSAMPE 731
 :|||||:
 Db 586 AAJTFNLGDLD---NYDEYDRSMLISQTSFERTFGLSTVHESTLMENGVPPDSYNPE 639

QY 732 NLIKEATHVISCYEDKSDWMEIGWIGYGVFEDLTGKFMHARGWHSIYCPKLPFKG 791
 :|||||:|||||:
 Db 640 TLIKEAIIHVISCYEEKTEWGKIGWIGYLSITEEDILGKFMHARGWHSIYCPKLPFKG 699

QY 792 SAPINLSDRLNQVRWALGSVELFSRCPICWYGYN-GRKLTLERFAYVNTIYPTISIP 850
 :|||||:|||||:
 Db 700 SAPINLSDRLHQWLRWALGSVELLSRCPICWYGYN-GRKLTLERFAYVNTIYPTISIP 759

QY 851 LLMVCTLLAVCLFTENQFTIPOQTSNIASTWLSLFLSIFATGILEMRWNSGVGIDBWRNEQ 910
 :|||||:|||||:
 Db 760 LWVACTLPAICLNGKFIPTLNSLMSFL-----GVSIEDLWRNEQ 802

QY 911 FWVIGGSAHLFAVFOQFLKMLAGDNTFTVSKAD-DLFEGELYTVKWTLLIPPSL 861
 :|||||:|||||:
 Db 862 LIINLVGWAQGSDALNGYEWGPLGKVFAFWTILHYPLFKGLMGRQRNTPTIVL 921

QY 1031 WSILLASIFSLIWNRIDPFTSRVTPGDPTECGIN 1065
 :|||||:
 Db 922 WSILLASIFSLIWNRIDPFTSRVTPGDPTECGIN 953

RESULT 10

T10800 cellulose synthase (EC 2.4.1.-) catalytic chain celA2 - upland cotton (fragment)
 C;Species: Gossypium hirsutum (upland cotton)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C;Accession: T10800
 R;Pear, J.R.; Kawagoe, Y.; Schreckengost, W.E.; Delmer D.P.; Stalker, D.M.
 Proc. Natl. Acad. Sci. U.S.A. 93, 12637-12642, 1996
 A;Title: Higher plants contain homologs of the bacterial celA genes encoding the catabolic enzymes involved in cellulose degradation
 A;Reference number: Z17152; MUID:97057296; PMID:8901635
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-685 <PEA>
 A;Cross-references:
 A;Experimental source: strain Acala SJ-2; fiber
 A;Genetics:
 A;Function:
 C;Description: involved in synthesis of cellulose
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 49.5%; Score 2808; DB 2; Length 685;
 Best Local Similarity 74.4%; Pred. No. 4.1e-214;
 Matches 512; Conservative 77; Mismatches 77; Indels 22; Gaps 6;

QY 397 RKWVPCFKYSTERPAPWYEAKKIDYLKDVKOTSEYKDRRAMKREEFKIRINALVSK 456
 :|||||:|||||:
 1 RRWVPCFKHNVDPRAPEFYNEKIDYLKDVKHPSFKERRAMKREEFKIRINALVSK 60

QY 457 AJKCPPECEWMODGTPWPGNNQDHGPMIQLVFLONGDAGNEELRLVYVSRKPGF 516
 :|||||:
 Db 61 AOKPPECEWMODGTPWPGNNQDHGPMIQLVFLONGDAGNEELRLVYVSRKPGF 516

QY 517 OHHKAGAMNALVRSAYLTNGPFILNLDCDIN 576
 :|||||:
 Db 121 OHHKAGAMNALVRSAYLTNGPFILNLDCDIN 576

QY 577 FRQFRDGTDKNDRYANRNTVERDINLGLDGQGPV 636
 :|||||:
 QY 577 FRQFRDGTDKNDRYANRNTVERDINLGLDGQGPV 636

Db 181 FPQRFGDIDRRHRYANRWWFFDINMLGLDGLOGPPVYVTCYCENRQALYGDPPVSKR 240
 Qy :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
 Db 637 KK-----PSLLSKLGSSRKNSKAKRE-----SDKKSGRH--TDSTVPEFL 678
 Db 241 PKMTCDCWSWCCCCGGSSRKSKKKGKGKGLLGLLYGKPKKMKGNVKGKGSAPYFDL 300
 Qy 679 DDFEGVSGAGFD-EKALLMSLNERKGOSAVFASSTLMENGVYPPSAIPNLKEA 737
 .301 EEEBEGLE--GYBLEKSTLMSOKNFEKRFQGQSPVFIASTLMENGGLPEGTNSTLSIKEA 358
 Qy 738 INVISCGYEDKSWDGMGECWGIKGSVTEILTGKMHARGWSRICMPKPAFKGSAPINL 797
 359 IIVISCGYEDKSWDGMGECWGIKGSVTEILTGKMHARGWSRICMPKPAFKGSAPINL 418
 Qy 798 SDRINQVIRWALGSVEILSFRHCPPIWYNGNRKFLERAYVNNTIYTPTISIPLMCTL 857
 Qy 419 SDRLHQVIRWALGSVEIFSRHCPPLWYGGKLUKLERAYNIVVFFTSIPLACTI 478
 Qy 858 LAVCLFLTNPQTIPQIPIQNSIATSIWTLSELSIFANGILEKWRWCGVIDEMWRNSOFWVIGV 917
 Db 479 PAVCLLTGKFIIPLNLSNLSWVHALFSLIATGVLELRWGSVSIQDWWRNEOFPWVIGV 538
 Qy 918 SAHLFAVFOGIKVLLAGIDTNEFTIKASDESDPFAELYLFKTTILLIPPTLILIVNG 977
 539 SAHLFAVFOGIKVLLAGIDTNEFTIKASDESDPFAELYLFKTTILLIPPTLILIVNG 597
 Qy 978 WAGVSYATNSGQSWGWLFGKUFAFWWVIVHXPFLKGIMGRNRTPTIVWWVSLIAS 1037
 Db 598 WAGVSYATNSGQSWGWLFGKUFAFWWVIVHXPFLKGIMGRNRTPTIVWWVSLIAS 657
 Qy 1038 IFSLWVWRDPPFPRSVRTDILCGINC 1065
 Db 658 IFSLWVWRDPPFPRSVRTDILCGINC 685

RESULT 11

Db 86157 hypothetical protein F22D16_26 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: D86157
 R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, D.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; anson, N.F.; Hughes, B.; Hultz, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.M.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shain, P.; Southwick, A.M.; Sun, H.; Tallon, A.; authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: D86157
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1181; <STOP>
 A:Cross-references: GB;AE005172; NID:96056428; PIDN:AAF02892.1; GSPDB:GN00141
 C:Genetics:
 C:Map position: 1

Query Match 40.9%; Score 2324; DB 2; Length 1181;
 Best Local Similarity 44.0%; Pred. No. 2,4e-175;
 Matches 491; Conservative 165; Mismatches 268; Indels 192; Gaps 29;

Qy 18 OTCOR-CSDNVGKTVDFDRFVACDICSPVCPCYEEVKERDKDENQSCPCKTYKRKGS 75
 Db 160 QICWLKGCD--RKVHG---RCE-CGRCRICKCYFDCITSGGNCPCGCKEPRDIND 211
 Qy 76 PAIPGDKDEGDLADEGTWEPNYQKEKISERMLGWLTRGKGEEMGEQYDKEYSHNHL 135
 Db 212 PETEEDDEDEEARP-----LHQ-----MGESKLKULS--WV 241

Qy 136 RLTSRQDQSFSAAASPERLSVSTIAGGKRLPYSSDVNQSPNRRTVDVPGLGNVAKER 195

Db 242 KSFKAQNQAGDFDHT-----RWLFETK-----GTYGKGNANPK- 275
 Db 196 VDGMKMKQEKNTPVSIQASERGGVTDASTDILADEALNDEAROPLRSRVSISPSRI 255
 Db 276 -DSTG-----GSGGGNGETPPE-----FBRSKRPLTRKYSVSAII 314
 Qy 256 NPYRMVIMRLVILCFLHYRITNPVNAFALMVSACEIWALSWILDQPPKFWVR 315
 Qy 315 SPYRLALRIVALVAGLFELTFWVRHPRNRAWMWGMGSTMICELWFALSMWLDQPLKLCVNR 374
 Db 316 ETYLDRLALRYDRBG--EP--SOLAADVSTVDPLKEPLPVANTVLSILAVDYPVD 370
 Db 375 LTDGVLKFERFSPNLNPKGRSDLPGIDVFNSTADPKEPLPVATVLSILAVDYPVD 434
 Qy 371 KVSCYVFDDGAMLSFEELAESTFARWVWPKKKSYSTEPRAPEWYPAKIQYDVKQVOT 430
 Db 435 KLACYLSDDGGALIIFELLAQATASFASTWVPEERKHNTEPRMEAYFGQKRNUFLNKYRL 494
 Qy 431 SFVKDRRAMKREVEEKIRINAL-----VSKALKC----- 460
 Db 495 DFRERRVRYKREYDEKFVIRINSPLRDRSADYDNVHIELRRAKKKOMEMMMGNPQENVI 554
 Qy 461 -PERGVWMODGTPWPG-----NTQDHPGMQIVFQGONG----GLDAEON----- 500
 Db 555 VPRAKTM-MSDGHSHWPWTNSGETDNRDHAQIQLQAMTAPPNAEPPVGAEEAENLDT 613
 Db 501 ---ELPRIVYVVERKREGFQHKKAGAMNALYRVSAVLTNEPFILNQDCHYINNSKAL 556
 Db 614 DVDTRLPMLVYVSEKRGYDHKKAGAMNALRTSAMSNGPFLNQDCHYINNSKAL 673
 Qy 557 REAMCFMDPNLKGQCVQVOPQRFGDKNDRYANRNTFEDINLRLGDIQGPVVGT 616
 Db 674 REGCMFLDGR-GDRICVYQFOPREGIDPNDRVAHNNTFEDVSMRALDGLQPMVGT 732
 Qy 617 GCVNRTALYGYEPPKVKHHKPSLSSLKUCGGSSRKNSKAKKESDKKKSGRHTDSTWVFE 676
 Db 733 GCIFRRTALYGFSPPRATEH-----GWGRKISLRRPKAMKKDEVSUPIN 783
 Qy 677 NLDDIEEGVEGASFDEDEKALLMSQMSLKEKGOSAVEAS-----TLMENG-- 722
 Db 784 GEENEENNDG---DINSSL---LPKRGFGNSNSFASIVPAEYQCRLIQDQGKCN 834
 Qy 723 -----GVPSPATPENLKEATHVTSQYEDKSDWMEIGWIGYVSETEIDLGEKHAR 775
 Db 835 SRPAGSLAVPREPLDAATVAEAISVISCYFEDKTWEKGKRVWGYSYTEDVWGYRMHR 894
 Qy 776 GWRSIYCMLPKPAFKGSAPINLSDRNQYLRWALGSVILFSRHCPIWGYNQRLKLER 835
 Db 895 GWRSIYCMLPKDAFGRATPMLNQYLDRHQLQYRWTGQSFESRNALP--ATTRMKLQR 952
 Qy 836 FAYVNTTYPITSPITMLMCTLAVCUTQNTIPQNSIATSTWFLSLSFATGMEL 895
 Db 953 VAYFNGMMPFTSFLIYLICILPAISLFSQFVQYQSDITFLYLSTLTCMLSLIEI 1012

Qy 896 RWGSGVGIDEMWRENEQFWVIGGVS AHLFAVFOGLIKLVLAGIDTNTFTVSKAS-DEDGD-F 952
 Db 1013 KWSGCTLHEWWRNNEQFWVIGGTSAHPAAVLQGLKLIKVAGVDISFTLTSKSSAPEDGDEF 1072

Qy 953 AELYLKFWTLLIPPTLILVNLVQVAGVSKVNSGQYOSWGLPLEGKLFKAFFWVWVLYP 1012

Db 1073 ADLWYKNSFSLMPPLPTIMMVNIAIAYGLARTYSPQWPSKLVGGVFFSWFLCHLYP 1132

Qy 1013 FLKGMLGRNRTPTIVWWVSLIASLISLWVRLD 1048

Db 1133 FAKGLMGRGRVTVFWSGLUSIISLWVYINP 1168

RESULT 12

T51546 cellulose synthase catalytic subunit-like protein - Arabidopsis thaliana
 N:Alternate names: protein F2K12_60
 C:Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
 C;accession: T51546
 R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mewes, H.W.; Mayer, A.; Reference number: Z25394
 A;Accession: T51546
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1145 <SAT>
 A;Cross references: EMBL:AL391141
 A;Experimental source: cultivar Columbia; BAC clone F2K13
 C;Genetics:
 A;Map position: 5
 A;Introns: 297/2; 566/3
 A;Note: F2K13_60

Query Match Similarity 40.6%; Score 2303.5; DB 2; Length 1145;
 Best Local Similarity 43.8%; Pred. No. 9.6e-174;
 Matches 490; Conservative 162; Mismatches 265; Indels 201; Gaps 28;

Qy 19 TCQI--CSDPVGKVTDGDRPVACDICSFPCRCYERYRKDGNGNQSCPQCKTRYKRLKGSP 76
 Db 127 SCAIPGCDAKVMSDERGQDILPCE--CDFKICRDCFIDAVTKGGTGGCPEKYK--NT 181

Qy 77 AIPEDKDEGLADEGTYEFPVNPQKEKISERMLGWLTRKGKEEMCEPQDKEVSHNLPR 136
 Db 182 HLTDQDVENG-----DQPRM---LPGGGSKM----- 205

Qy 137 LTSRQDTSGEFSASAESPRL-VSSTIAGKRLPYSSDVNVSPNRRIVDPV--GLGNVAW 192
 Db 206 -----ERRISMVKSTINKSALMRSGTGDFH - NRWFLETFGTGTYGNAFW 248

Qy 193 KERVGWKMKOKEKNITGPVPSITOQAASERGKGDIDASTDILADEAL-LNDEAROPLRSKRVSIP 251
 Db 249 -----TDKGDFGSKGDKGDDGGMGMEAODMSRPMWRPLTRKLKP 289

Qy 252 SSRINPVPYRMVIMLRUWVOLFLHYRITNPVNPAFLWLVSVICETWFLSWSLWDFOPKWF 311
 Db 290 AGVISPYRLIFITRIVLWVFLTRVKHQNDPAVWLMGSVCLWFAWSLWLUOPKLIC 349

Qy 312 PVRRETYLDRLAQLRD--REGEPE--SQLAVIDFVSTWDPLKEPPLVANTVISLAVD 366
 Db 350 PINRATDQLVLERKFETPTPASNPITGKSDLPGFDFVSTADPEKEPPLVANTISILAE 409

Qy 367 YPVDRKVSTCYFDDGAAMLSPESLAETSEPARWKWPFCKKYSTEPRAPENYFAAKIDYKD 426
 Db 410 YPVERKLSCIVSDDGALLTETEAMERAASFANIWVFPCRKHAIERPNPDSYFLSKRDPYKN 469

Qy 427 KVQPSFKVORRAMEKREYEFKIRINALVS----- 456
 Db 470 KVKSDFVKDRRVEKFDEFKVRVNSLPPDSIRRNSDAYAHAREEIKAMKMQQRQNDRDEPME 529

Qy 457 ALKCPPEEGWVMQDCTPWPW-----NWTGDHGMQVFLG-----QNGGLDAE 498
 Db 530 PVKIPKATW-MADGTHWPGTWTLSASDHAKGDHAGIIQMLKPPSPDEPLPHGVSGFLDT 588

Qy 499 GNE--LPRLYVVSKEKPGFQHKKAGAMALVRYSAVIITNGPFTINDCDHYINNSKAL 556
 Db 589 DVDIRLPLLVYVSREKPOYDHNNKAGAMALVLRAASIMSNGPFTILNCDHYTINSEAL 648

Qy 557 REAMCFLMDPLGKQCVYOFPOEDGIDKNDRYANTRNFEDINLRGLDGQGQPVYGT 616
 Db 649 REGCMCFMDRG-GDRCLYVQFOPREGTIDPSDRANHNTFEDNMRALDGLMPVYVWGT 707

Qy 617 GCVNRTALYGYEPPIVKHKPKPSLLSKLCGGSSKKN---SKAKKESKKKSRHTDST 672
 Db 708 GCLFRRLALYGFNPB-RSKDFSPCWCSCCPRSSKKKNTPEENRALMSD----- 755

Qy 718 VPVFNLDIEEGVEGAGFDEDEKALMSQSLEKRFGQSAVFVAS-----TL 718

Db 756 -----YDDEE--MNLSLVPKRFQNSPLFIDSLIPVAFQGRPLADHRA 795

Qy 719 MENGVPPSAT-PENL---KEAIHV1SCGYEDKSDMGMEIGWIYGSVTEDILTGFEM 772

RESULT 13
 T05646 hypothetical protein F20D10_310 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse ear cress)
 C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
 C;Accession: T05646
 R;Bevan, M.; Wedder, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, A.; Reference number: Z15420
 A;Accession: T05646
 A;Molecule type: DNA
 A;Cross references: EMBL:AL035538
 A;Experimental source: cultivar Columbia; BAC clone F20D10
 C;Genetics:
 A;Map Position: 4
 A;Introns: 139/2; 675/3
 A;Note: F20D10_310

Query Match Similarity 40.3%; Score 2289.5; DB 2; Length 1111;
 Best Local Similarity 43.1%; Pred. No. 1.2e-172;
 Matches 481; Conservative 158; Mismatches 256; Indels 221; Gaps 28;

Qy 23 CSDNYGKTVDGDRPVACDICSFPCRCYERYRKDGNGNQSCPQCKTRYKRLKGSP 82
 Db 120 CDGNYNKDERGKDVMPCE-CRFKICRDCFCMDAQKE-TGLCPGCKEQYK-----IGDL 169

Qy 83 DEDGL--ADEGTYEFPVNPQKEKISERMLGWLTRKGKEEMCEPQDKEVSHNLPLTSQ 141
 Db 170 DDFTDPSGSGALPLAPKGQ-----RGNNMMS-----MMKR 202

Qy 142 DTSGERSAASPERLVSSTIAGKRLPYSSDVNVSPNRRIVD--PVGLGNVAW-KERVD 197
 Db 203 NQNGEE-----DHRWFLETFGTGTYGNAFWQDEM 234

Qy 198 GWKMKQEKNTGPVPSITOQAASERGKGDIDASTDILADEALLNDEAROPLRSKRVSISRSRNP 257

Db 235 GDDMDEGMRRGMVETA----- 300
 Qy 258 YRMVIMLRLVYLICLHYRITNPVNPAFLWLVSVICETWFLSWSLWDFOPKWPVNR 317
 Db 271 YRLRIVLTFVFLCFLPFLTRNPNSDIAWILMISCELUMFGFSWLDQPLCPINRST 330

Qy 318 YLDRJALRYDREGEPE----SQLAVIDFVSTWDPLKEPPLVTNLVSLAVDYPVDR 372
 Db 331 DLEVLRKDFDMPSPNPTGRSLDGLPFIDLVFSTADPEKEPPLVTNLVSLAVDYPVDR 390

Qy 373 SCYVFDGGAMLSFSLAETSEFARKWVPECKKSYTEPRAPEWVFAAKIDYKLKQTSF 432

Db	391	SCYLSDDGALLSFEAMAEASFADLWVPPFCRKHNIEPRNPDSYFLSKIDPTKNKSRDF	450
Qy	433	VKRRRMRREYEFKIRINALVS-----KAL-----	463
Db	451	VKDRRKIKREYDEFKVRINGLPDSIRRSDAFAEEMKALKQMRPESGGDPTERVKPKA	510
Qy	464	GWMQDGTPWPG-----NNIGDHPMIQVFLGQNGGLDAEGN-----EL	502
Db	511	TW-MADGTHWPGTWAASREHSGKDAGILQMLKKPSSDPLIGNSDDKVIFSDTDRL	569
Qy	503	PRUVYVSZEKRCFOFHKGAKAGANALVRVSAVLTNGPFTNLNCQDHYNNSKALREANC	562
Db	570	PMFVVYVSREKRPGYDHNKAGAMNALVRASALTSNGPFTNLNCQDHYTYNCKAVREGM	629
Qy	563	LMDPNLGKOVYQFQRFDGIDKNDRYANRNTVFDINLRGLDQGPVYGTGVNF	622
Db	630	MMDRG-GRDICYIQFQREGIDPSDRVANNNTVFGDNRMALDGVQCPVYQGTGW	688
Qy	623	TALYGYEPPIVKVHKPSLISKLCUGGSRKNSKAKKEDDKKSGRHTDSTPVFNLDIE	682
Db	689	FALIGFDFP-----NPDKLLE-----KKESE-----T	710
Qy	683	EGVEGAGIDDEKALLMSOMSLERFQGS-----AVFVASTLMENGGV---PSA-	728
Db	711	EALTTSPDTPD-----LDTVOLPKRFGNSTLLEASPIPAFQGRPLADHPAVKGRPGC	766
Qy	729	-TPENLL-----KEAHIVISCGEDKSDMGMEWTGWWIGSVTEDILTGFKHARGWRSY	782
Db	767	RVRDPDIDATVAESVSYISCVWEDKTEWGDRGYWIGSVTERVTGVRMHNGWRSY	826
Qy	783	MPLKPAFKGSAPINLSDRDLNQVRLWALGSVETLFSRHCPIWGYGNGRKFLERFAYNT	842
Db	827	ITKRDSFRGSAPINLTDRLHQVLWATGSVEIFFSRNNAIL--ASKRLKFLQLAYLN	884
Qy	843	IPITSIPLIMCTTLAVCLFLNQFIIQPNISIASIWIWLSLFSIFATGILERWMSGV	902
Db	885	IYPTSFLLFLYLICFLPAFLSFLQGIFTVLTLSFLVLYMITICLIGAVLEYKWSG	944
Qy	903	DEWRNEOFWVIGGSVHSAHLFAYFOGLKVLLAGIDTNFTVTSKASDED- -FAELYLK	960
Db	945	EEWNRNEOWMIALGTSISSLVYAVQGVLKVIAGIEISFLITKSGDDNEDIVADLY	1004
Qy	961	TWLIPIPTTLTVNLGVWAGVSYATANGYQSWGPPGKLEFFAWFVIVHYPPLKGLM	1020
Db	1005	SSLMIPPIVIVAMNITIAIWAFRTTYQAVPOMSKLIGGAFSSFWLHALHYPEAKGL	1064
Qy	1021	QNRTPTIIVVWSLLASIFSLWVRIIDPFTSRVTP 1056	
Db	1065	RKTPTIVEVWAGLIATTISLWTAIMP---NTGP 1096	
RESULT 14			
Db	d84741	probable cellulose synthase [imported] - Arabidopsis thaliana	
C;Species:	Arabidopsis thaliana (mouse-ear cress)		
C;Date:	02-Feb-2001 #sequence_revision 02-Feb-2001		
C;Accession:	D84741		
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Weiss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Wenter, J.	Nature 402, 761-768, 1999		
A;Title:	Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.		
A;Reference number:	A84420; MUID:20083487; PMID:10617197		
A;Accession:	D84741		
A;Status:	Preliminary		
A;Molecule type:	DNA		
A;Residues:	1-1036 <stop>		
A;Cross-references:	GB:AE002093; NID:92924781; PIDN:AAC04910.1; GSPDB:GN00139		
A;Gene:	At2g33100		
A;Map position:	2		
Best local similarity 45.0%; Pred. No. 2_6e-167; Matches 469; Conservative 150; Mismatches 248; Indels 175; Gaps 26;			
Matches	469;		
Conservative	150;		
Mismatches	248;		
Indels	175;		
Gaps	26;		
Qy	117	GEEMCEPOQDKEVSHNPLRTSROTSCEFSAASPERLUSVSTIAGG-----KR	166
Db	48	GDKSGONDY---INTVLMPTPFDNQPGSSGTSISESKGDANRGGGGSGPKMKLERR	104
Qy	167	LPYSSDVONS-----PNRRTIVD---PGVGLGNYAWKERYDGWAKMOKRKTGPVST	212
Db	105	LSVMKSNKSMRSLRSQSGPFDHNRKWLLESKGKGWAFNNSPDDYD-----	152
Qy	213	QASERGGDIDASDILADEALLEARQPLSKRSVSPSSRNPKVSYRVMRVLICF	272
Db	153	-----GGV-----SKSDFL-----DKPHKPLTRKVQPAKILSPYRLLIVRIVVF	196
Qy	273	LYRILTPVNPNAFLWLVSYVICETWFALSWILDOPFKWPVNRETYLDRLAYDREEP	332
Db	197	LWWRITNPNEADAMLWLGJSIVCEIWFASWILDILPKLNPINRAT--DLVALHDKFEQ	253
Qy	333	-----SQLAADIFVSTVDPLKEPLVANTVLSIADVDPVDKSYWVDDGA	394
Db	254	SPSNPTGRSDLPGLGVDFVSTADEPEKEPLVANTLTSIADVDPKIEKUSLDDGAI	313
Qy	385	SFESIAETSEPARKWPVERKYSSEPRADWYFAKTDYKLKVOTKQVTDYKRVDRAMKREY	444
Db	314	TFEAMAEAVRAFAEWVWPFCRKHDIEPRNPDSYFSIKOPTKKRQDFVKDRWKIREYD	373
Qy	445	EFKRINALVSKA-----CPEG-----WMMQDCPW	473
Db	374	EFKVRINGLPDEOIKKRAEOFNMRELKEKRAKEBKGKVLPPDGQEVVYKATW-MADGHW	432
Qy	474	PG-----NNIGDHPMIQVFL-----GQN-GGLDAEGNEL--PRLVYVSRE	512
Db	433	PGTWEFPKPDHSKCDHAGILQIMSKVPDLEPVWGMGPNEGALDTGIDRVPMAYVSREK	492
Qy	513	RGFOHHRKAGAMNALVRVSAVLTNGPFTNLQDCHYNTNSKALREANCFLMPNGLR	572
Db	493	RPGFDHNNKKAGAMGMVRASALTSNGAFLNLQDCHYTYNSKAIKEGCFMMDRG-GDRI	551
Qy	573	CYQFOPREDGIDKNDRYANRNTVFDINLRGLDQGPVYGTGVCNTRALYGYEPI	632
Db	552	CYTOFPQRPEGIDPSDRVANNTVFDGMRALDGLQDPVYVGCMFRYALYGFNP-	610
Qy	633	VKHKHKPSLISKLCUGGSRKNSKAKKEDDKKSGRHTDSTPVFNLDIEGVEGAGFDD	692
Db	611	-----RANEYGVFGQEKAPAMHRTQSOSQTS-----ASDLESITQPLNDP	655
Qy	693	EKALIMSMOSLERFQGOSAV-----VASTLMENGGVPPA-----TPEN	732
Db	656	-----LGPKKFGNSTMTDTPVAYEQGRPLADHMSVKRNGRPPGALLRPPLDAP	708
Qy	733	LKEAHIVISCGEDKSDMGMEWGIWIKYGSVTEILTGFKMHARGWRSYTCMPLPAFKGS	792
Db	709	TVAEAIAYVSCWEDNTFWDGDRIGWIGSYTDEVVTGVRMHNRGWRSTCYTKRDAFRGT	768
Qy	793	APINLSDRDLNQVRLWALGSVETLFSRHCPIWGYGNGRKFLERFAYNTTIVITSIDL	852
Db	769	APINLTDRLHQVLWATGSVEIFFSKNNAMF-ATTRRKLFLQRAYLNIVGIVFTSIFL	826
Qy	853	MCYCLLAVCLFLNQFIIQPNISIASIWIWLSLFSIFAT---GILEMWGSGIDEMRN	908
Db	827	YCFPLALCGLFSLKFIYVOSL-----DIHFLSYLICITVTLTLLSILVERWGSGLGEEMRN	882
Qy	909	EQFWVIGGSVHSAHLFAYFOGLKVLAGIDTNFTVTSKASDED- -FAELYLKWTILP	966
Db	883	EQFLIGGTSASHAALWVQGLKVIAGIEISFTUTSKASGEDDIDFADLYTKWTGLFIM	942
Qy	967	PTTLLIVNLGVWVAGVSYATANGYQSWGPPGKLEFFWVTHYPFLKGMRQRNRPRT	1026
Db	943	PLTIIIVNLVATVIGASRTIYSVIPOMGKLMGGIFFSLWVLTWVHYPFLKGMRQRKVT	1002

RESULT 15
C86446 C86446 probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Accession: C86446 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Fedderspiel, N.A.; Kaul, S.; White, O.; Devayen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dews, A.N.; Fries, N.F.; Hughes, B.K.; Hulzar, L.; Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khavkin, E.; Rizzo, C.A.; Li, Y.; Li, X.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marr, R.; Rooney, T.; Rowley, D.; Sakano, H.; Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Toker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86446
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-979 <STOP>
A;Cross-references: GB:AE005172; NID:910801364; PIDN:AAG23436.1; GSPDB:GN00141
A;Map position: 1
C;Genetics:
Query Match Best Local Similarity Score DB length Gaps Matches
Qy 126 DKEFVSHNLPLRLTSRQDTSGEFSASA PERLSVSSTIAGKRLPYSDVNOSPNRRIYDPV 38.5% 2185.5; DB 2; 979; 247; Indels 151; Gaps 25; 45.5%; Pred. No. 1.7e-164; 463; Conservative 157; Mismatches 186 GLGNVANKWERDQWKKM**Q**EKTNTGCPVSTQASERGGVTDAS-TDIL-----ADEALIN 48 SLSIVILLE-ID---SNOESPVSVSGDIVSGSSKDNEPDLTDRINGEEFEEDTILS 238 DEARQPISKVSPSRNPRYAVIMLRLVTCFLFHRYITNPVPA[PALW]YVICEW 102 -KISYSLJRVKISPII[ALYV]LIRVVLWVSLAFIRRNINKALWLWLSVICELW 297 298 FALSWILDQFPKWPFWPNRETYLDLALRY---DREGEPSQLAANDVIFVSTDPLKBP 161 FAFTSWLQDQIKPKPWNHATDLEAKAFETPNPDNPNGKSDPGLDVFVSTDAAEKPP 353 LVTANTVSI[LAVD]YVPUKVSCVFDGCAAML[FESLAETSEFARKWVFPCKY]SIERA 221 LYTAINTLISLSDVYPVKEKLSVYISDDGGSLV[FETAAEAASFAKI]WVPCFKRKHKIEPRN 413 PEWYFAAKIDYDKQVQSFVKRRAKMRREYEEFKIRNALV-----SK---- 281 PESYFGKLRDPYKVRIIDFVRERRVYRAYDEPKVRYNPHSIRRSDAFNSKEEKA 456 457 -----ALKCPERGWVMODGTPWPG-----NNTGDHPGMTOV 341 LEKWKHMKVYKVEEQDQIKEPRPAVLPKA[TW-MSDGTHRGPTWAVSGPHRSRGDHASVIQ 488 FL-----G QONGGLDAEGNE --LPRLYVVSREKRPQFOHQHKAGAMMALVRYSA 533 400 LLDPGPDPVEKGKGEGRALDLEGVDIPLMLYVSVREKRPQGDHKKAGAMMALVRYSA 534 VLTNGPFTLNLOCDHDYTNNSKALREAMCFLMDPDLNGKOCVYQFQREDFGIDKNDRYAN 594 NTYFFDIDLRLGDCIGQGVYVWVGTGCCVNRTALYQEP -IKVHKHKPSLISKLOGSRKK 519 NTYFFDIDLRLADLGIOGPWVGTGCLFRRTALYQFNPDVWEEPS -GSTCFPLIKK 653 NSIAKKEKDKKSGRHTDSTVPFNLDDIEEGVEGAGFDEDEKALLMSMSLEKRGOSAV 460 IMNGPPEFLNLDCHDYYVNSRAFRDGCMMFD -GDGRSYVQFQRREGIDESDRYAN 518 594 NTYFFDIDLRLGDCIGQGVYVWVGTGCCVNRTALYQEP -IKVHKHKPSLISKLOGSRKK 519 NTYFFDIDLRLADLGIOGPWVGTGCLFRRTALYQFNPDVWEEPS -GSTCFPLIKK 576 653 NSIAKKEKDKKSGRHTDSTVPFNLDDIEEGVEGAGFDEDEKALLMSMSLEKRGOSAV 712

Search completed: June 16, 2003, 11:08:32

THIS PAGE BLANK (USPTO)

Db 181 IVPVGLGNWAKWERVDGPKMKQDKBKTGCVSTQASERGGVDDASTDILADELLNDEA 240 ; PRIORITY APPLICATION NUMBER: 09/720383
 QY 241 RQPLSRKVSSRSPSRNPYRMVIMRLVLTICLFLHYRITPVPNAFLWIVSVCIEIWAL 300 ; PRIOR FILING DATE: 2000-12-21
 Db 241 RQPLSRKVSSRSPSRNPYRMVIMRLVLTICLFLHYRITPVPNAFLWIVSVCIEIWAL 300 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO: 30
 ; LENGTH: 1080
 ; TYPE: PRT
 ; ORGANISM: *Triticum aestivum*
 ; US-09-000-237-30
 Query Match 79.4%; Score 4507; DB 10; Length 1080;
 Best Local Similarity 78.4%; Pred. NO. 0; Mismatches 107; Indels 26; Gaps 8;
 Matches 836; Conservative 98; MisMatches 107; Index 26; Gaps 8;
 QY 421 IDVYKDKVOTSFVKDRRANKREVEEFKIRINALVSKALKCPEEGWMODGTPMPEGNNITGD 480 ;
 Db 421 IDVYKDKVOTSFVKDRRANKREVEEFKIRINALVSKALKCPEEGWMODGTPMPEGNNITGD 480 ;
 ;
 QY 481 HPGMLOVFGONGSLDARECNELPPLVYVSREKRGFOHKRKAGAMNALVRVSAVLTINGF 420 ;
 Db 481 HPGMLOVFGONGSLDARECNELPPLVYVSREKRGFOHKRKAGAMNALVRVSAVLTINGF 420 ;
 ;
 QY 541 ILNLDCDHYNNSKALREMCFLMDPNLGKQCVYQFOPRFDGIDKNDRYANRNTVFDI 600 ;
 Db 541 ILNLDCDHYNNSKALREMCFLMDPNLGKQCVYQFOPRFDGIDKNDRYANRNTVFDI 600 ;
 ;
 QY 601 NLRLDQGQGPVYGTGCVNRALYGEPEPIKKHKKPSLSSKLCGGSSRRKKSKAKES 660 ;
 Db 601 NLRLDQGQGPVYGTGCVNRALYGEPEPIKKHKKPSLSSKLCGGSSRRKKSKAKES 660 ;
 ;
 QY 661 DKKSGSRHDSTVFWNLDDIEVEGAGFDEKALLMSOMSLERKGOSAVFWASTME 720 ;
 Db 661 DKKSGSRHDSTVFWNLDDIEVEGAGFDEKALLMSOMSLERKGOSAVFWASTME 720 ;
 ;
 QY 721 NGGVPPSATPENLKEAHTHVISGYEDDSDWGMIGWTGSVTDLGFKMPARGWISI 780 ;
 Db 721 NGGVPPSATPENLKEAHTHVISGYEDDSDWGMIGWTGSVTDLGFKMPARGWISI 780 ;
 ;
 QY 781 YCMPLKPAFKGASPAPINSLDRNQVLRLWALGSVELLSRCPINWYNGNRLKFLERAYN 840 ;
 Db 781 YCMPLKPAFKGASPAPINSLDRNQVLRLWALGSVELLSRCPINWYNGNRLKFLERAYN 840 ;
 ;
 QY 841 TTIVPITSPLIMCILAVCLTQNQFIPOISNIASWFLSFLSITAGILEMRWAG 900 ;
 Db 841 TTIVPITSPLIMCILAVCLTQNQFIPOISNIASWFLSFLSITAGILEMRWAG 900 ;
 ;
 QY 901 GIDEMWRNQFWVIGVSHLRAVFOGILKVLAGIDNTVTISKASDDEDFEALYLRKW 960 ;
 Db 901 GIDEMWRNQFWVIGVSHLRAVFOGILKVLAGIDNTVTISKASDDEDFEALYLRKW 960 ;
 ;
 QY 961 TTLIIPPTPLTILVLUVGWAGVSAINSQYQSMPLFGKLFEEFWIYVLYPFLKLMGR 1020 ;
 Db 961 TTLIIPPTPLTILVLUVGWAGVSAINSQYQSMPLFGKLFEEFWIYVLYPFLKLMGR 1020 ;
 ;
 QY 1021 QNRPTIVWWSVLAESTSLWLRIDFTSRVGPDLIECGNC 1065 ;
 Db 1021 QNRPTIVWWSVLAESTSLWLRIDFTSRVGPDLIECGNC 1065 ;
 ;
 RESULT 2
 US 09 900-237-30
 ; Sequence 30, Application US/09900237
 ; Patent No. US20020120124A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Stephen
 ; TITLE OF INVENTION: Plant Cellulose Synthases
 ; FILE REFERENCE: BBL170 US CIP CURRENT APPLICATION NUMBER: US/09/900, 237
 ; CURRENT FILING DATE: 2001-07-06
 ; PRIOR APPLICATION NUMBER: 60/092, 844
 ; PRIOR FILING DATE: 1998-07-14
 ; PCT APPLICATION NUMBER: PCT/US99/15871
 ; PRIORITY FILING DATE: 1999-07-13
 ;
 Db 674 SSDKRSRHTDSTPVFNLDDEEGVAGFDEKALLMSOMSLERKGOSAVFWASTL 733 ;
 Qy 719 MENGGVPPSATPENLKEAHTHVISGYEDDSDWGMIGWTGSVTDLGFKMHARGWR 778 ;
 Db 734 MEYGGVPOSSTPSLILKERAHTHVISGYEDDSWENGETEICWIYGSVTEDTLGFKMHARGWR 793 ;
 ;
 Qy 779 SIVCMPLKPAFKGASPAPINSLDRNQVLRLWALGSVELLSRCPINWYNGNRLKFLERAY 838 ;
 Db 794 SVVCMPLKPAFKGASPAPINSLDRNQVLRLWALGSVELLSRCPINWYNGNRLKFLERAY 853 ;
 Qy 839 VNTTIVPITSPLIMCILAVCLTQNQFIPOISNIASWFLSFLSITAGILEMRWWS 898 ;
 Db 854 INTTIVPITSPLIMCILAVCLTQNQFIPOISNIASWFLSFLSITAGILEMRWWS 913 ;

QY 899 GVGIDEMWRNEQFWVIGGSVAHLFAVFOGILKVLAGIDTNTFTSKASDEGDFAELYLF 958
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 914 GVGIDEMWRNEQFWVIGGSVAHLFAVFOGLKLVLAGIDTNTFTSKASDEGDFAELYF 973
 QY 959 KWTTLILPPTLILIVNLUWVWAGVSTAINSGYQSMGPGLGALFEAWVIVLYPELKLM 1018
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 974 KWTTLILPPTLILIVNLUWVWAGVSTAINSGYQSMGPGLGALFEAWVIVLYPELKLM 1033
 QY 1019 GRQRNRPITIVWWMSVLLASIFSLWLRIDPFTSRVIGPDILECGINC 1065
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 1034 GRQRNRPITIVWWMSVLLASIFSLWLRIDPFTSRVIGPDILECGINC 1080
 RESULT 3
 US-00-838-539-8
 ; Sequence 8, Application US/09838539
 ; GENERAL INFORMATION:
 ; APPLICANT: Stalker, D. et al.
 ; PATENT NO. US20020129401A1
 ; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
 ; TITLE OF INVENTION: Sequences
 ; FILE REFERENCE: 15621/03/US
 ; CURRENT APPLICATION NUMBER: US/09/838,539
 ; CURRENT FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/029, 987
 ; PRIOR FILING DATE: 1996-10-29
 ; PRIOR APPLICATION NUMBER: 08/960, 048
 ; PRIOR FILING DATE: 1997-10-29
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 881
 ; TYPE: PRT
 ; ORGANISM: Oryzae sativa
 US-09-838-539-8
 Query Match 72.8%; Score 4133.5; DB 10; Length 881;
 Best Local Similarity 86.8%; Pred. No. 0;
 Matches 766; Conservative 55; Mismatches 56; Indels 5; Gaps 3;
 QY 188 GNVAWKBRVDGSKMKQBKKNTGPV--STQASERGGY-DIADASTDILADEALNDERRQP 243
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 1 GNVAWKBRVDGSKMKQBKKNTGPV--STQASERGGY-DIADASTDILADEALNDERRQP 60
 Db 244 LSRKVSPSSRINPYRMVIMURLVICLFLHYRITNPVNAFLMLVSYICEIWALSWI 303
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 61 LSRKVSPSSRINPYRMVIMURLVICLFLHYRITNPVNAFLMLVSYICEIWALSWI 120
 ; SEQ ID NO 26
 ; LENGTH: 1091
 ; TYPE: PRT
 ; ORGANISM: Impatiens balsamia
 US-09-900-237-26
 Query Match 69.5%; Score 3999.5; DB 10; Length 1091;
 Best Local Similarity 69.5%; Pred. No. 0;
 Matches 750; Conservative 126; Mismatches 162; Indels 41; Gaps 14;
 QY 364 AVDPYDKVSCVVFDDGAMLFSESLAETSEPARKVNPFCKYSIRPAPENWYAAKIDY 423
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 181 AVDPYDKVSCVVFDDGAMLFSESLAETSEPARKVNPFCKYSIRPAPENWYAAKIDY 240
 Db 424 LKDVKYPSFVDRRANKREYEFKIRINALYNSKALKCPEEGKWMQGCTPWPGNNNGDHPG 483
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 241 LKDVKYPSFVDRRANKREYEFKIRINALYNSKALKCPEEGKWMQGCTPWPGNNNGDHPG 300
 ; SEQ ID NO 26
 ; LENGTH: 1091
 ; TYPE: PRT
 ; ORGANISM: Impatiens balsamia
 US-09-900-237-26
 Query Match 70.5%; Score 3999.5; DB 10; Length 1091;
 Best Local Similarity 69.5%; Pred. No. 0;
 Matches 750; Conservative 126; Mismatches 162; Indels 41; Gaps 14;
 QY 484 MTQVFLCQNGGLADEGNEPLRVVSYREKRGCFOHKKAGAMNALYRSAVLINGPFLN 543
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 301 MTOVFLGHSGGDTECNELPRIVVSREKRGCFOHKKAGAMNALYRSAVLINGPMLN 360
 Qy 544 LDCDHYTKINSKALREAMCFLMDPNLGKQCVYOFPRFGDIDKNDRYANRNTFFDLN 603
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 361 LDCDHYTKINSKALREAMCFLMDPNLGKQCVYOFPRFGDIDKNDRYANRNTFFDLN 420
 Qy 604 GLDGIGOPWVYTGCVFNRITALYGYRPPKIKHKKSLSKLGGGRKRNKAKKESDKK 663
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 421 GLDGIGOPWVYTGCVFNRITALYGYRPPKIKHKKSLSKLGGGRKRNKAKKESDKK 479
 RESULT 4
 US-09-900-237-26
 ; Sequence 26, Application US/09900237
 ; Patent No. US2002012012A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALLEN, Stephen
 ; TITLE OF INVENTION: Plant Cellulose Synthases
 ; FILE REFERENCE: BB1170 US CIP
 ; CURRENT APPLICATION NUMBER: US/09/900, 237
 ; CURRENT FILING DATE: 2001-07-06
 ; PRIOR APPLICATION NUMBER: 60/092, 844
 ; PRIOR FILING DATE: 1996-07-14
 ; PRIOR APPLICATION NUMBER: PCT/US9/15871
 ; PRIOR FILING DATE: 1999-07-13
 ; PRIOR APPLICATION NUMBER: 09/720383
 ; PRIOR FILING DATE: 2000-12-21
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 26
 ; LENGTH: 1091
 ; TYPE: PRT
 ; ORGANISM: Impatiens balsamia
 US-09-900-237-26
 Query Match 70.5%; Score 3999.5; DB 10; Length 1091;
 Best Local Similarity 69.5%; Pred. No. 0;
 Matches 750; Conservative 126; Mismatches 162; Indels 41; Gaps 14;
 QY 10 KPKMKNIWPQTQCISDNVKGKTVDGDRWACDCISFPVCPRCYERYERKGQNSCPCQKTRY 69
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 91 KFQKGSPRVEGDEEDV-DDLENEFNYKGKKNOK---VTARRPWQD-QDIEL 143
 Qy 130 S---RHH---LPRLTTSRDTGERSAASPERLSSTSING---GKRUPYSSDVNQ 176
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 70 KRLKJGSAIPGKDDGGLADEGVFENPQKEKISERMGLWHTRGKGEEMGEPQDKEV 129
 ; SEQ ID NO 26
 Db 144 SVSSRSHDESQQPVPLJLGHHSVGET-PTPDNHSLRTSGPPIGPWEKSIYIDPRQPV 201
 Qy 177 PRRIWDP----VGLGNVAKERVQCKMOKEDKNGTPVSTQASERGGDIDASDIL 230
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 202 AYRITIVPSKDSLNSYGINGVDRKERVGWLQKEKNQMSRYPCKG-DTEGPGSNG 259

QY 231 AREALLUNDEARQPLSRKVSISSRINPYRMIVMLVTCILHYRTNPVNFAFLWV 290 ; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
Db : ;
; SEQ ID NO: 4
; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Zea mays
; US-09-900-237-4
Query Match 69.1%; Score 3922; DB 10; Length 1148;
Best Local Similarity 67.2%; Pred. No: 0; Matches 734; Conservative 141; Mismatches 161; Index 56; Gaps 17;
Matches 734; Conservatve 141; Mismatches 161; Indexs 56; Gaps 17;
Db 320 SVICEIWFAWSWILDQOPPKWPVNREYDRLALRYDREGEPSOLAANDIFSTVDLKE 379
QY 351 PPLVTANTVWLSTIAVDPVKVSCYVFDDGAAMLSFESLAETSEFARKWVPCCKYSIEP 410
Db 380 PPLYANTVWLISLAVDPVKVSCYVSDGSAMLTIFERSETAEFAKKWAPCKKHSIEP 439
QY 411 RAPEWYFAAKIDYLKDQVTSFKDRRAMKREVEEKFTRINALVSKALKCPEGWMDG 470
Db 440 RAPEFYFAQKIDYLKDQVSPFVKERRAKREVEEKFVIRNALVAKAQVPEEGWTMDG 499
Qy 471 TPRPGNNTGDHPCMIVFLGONGQNLDAEGNEELPRLVVSREKRGFOHHKAGAMNAVR 530
Db 500 TPRPGNNSRDHPGMIQFLGHSGGDFREGNELPRLVVSREKRGFOHHKAGAMNALR 559
Qy 531 VSAVLTINGPFILNLDCDIYINNSKALRATMCFLMDPNIGKQCYVQFQRFQDGIDKNDY 590
Db 560 VSAVLTINGAYLLNVDCHYFNNSKCLKEACMFMDPNUKGKTCVQFQRFQDGIDLADY 619
Qy 591 ANRNTVFFDINFRLGLDQGPGVYTGSCFVNRYALGYEPPKVKHCRPLSKLGCSR 650
Db 620 ANRNVIFEDINLKGLDGQGPVYTGCCNRQALYGDVPILEPTILIKSCGSR 679
Qy 651 KKNASKA-KKEDKKSSGRHTDSTPVNLDDLEEGVEGAFDEKALLMSQMSLEKFGQ 709
Db 680 KKGKGGGNKYYDKRNRAKRTTESTAPIFNMEDIEGIE--GYDERSFLMAO-SYEKRGQ 736
Qy 710 SAVVFVASTLMEENGVPSPSATPENLKEAHVSCGYEDKSDMOMEIGMIGYSYTEDILT 769
Db 737 SPVLIATTEMQQGLPSTINSATLKEAHVSCGYEDKTEKNGEIGWIGSYTEDILT 796
Qy 770 FKHARGHRSTICMPKLAFKSAPINLSDRNLNOLVRLAWGSEVILSRSRCIHWYGNGR 829
Db 797 FKHTTRGNISITCMPMPRAFKGSAPINLSDRNLNOLVRLAWGSEVILSRSRCIHWYGSGR 856
Qy 830 LKLERFVWNTIYPTISIPLIMCPLAVCLFTNOLIPQISNISIWFSLFSLIFA 889
Db 857 LKFLERLAYINTIVYPLISIPLIACLTPLAICLTGKTFVPEISNYASIWFLFVIFS 916
Qy 890 TGLEMRSVGVIDEMWRNEQFWIGGVSAHLJAVFOGILKVIAGIDNTFTTAKSDED 949
Db 917 TGFLERLNGSVTLEDWNRNEQFAVIGGSAHLJAVFOGILKVIAGIDNTFTTAKSDED 976
Qy 950 GDDAELYKFWTMLIIPPTTLVNLQVAGVAGVSYAANSGYOSWGPJLGFKAFFWIVH 1009
Db 977 GDPAELYKFWVLSLIPPTTLVNMVGIVAGVFSAINSGYOSWGPJLGFKAFFWIVH 1036
Qy 1010 LYPFLKGIGMRONRPTLVWMSVLSASIFSLWLRDPFTS--RVTGPDILEGGINC 1065
Db 1037 LYPFLKGIGRQNRTPTIVVWMSVLLASIFSLWLRDPFTSDSTKARG---OCGIDC 1091
RESULT 5
US-09-900-237-4
Sequence 4, Application US/09900237
; Parent No. US20020120124A1
; Applicant: Allen, Stephen
; Title of Invention: Plant Cellulose Synthases
; File Reference: BB1170 US CIP
; Current Application Number: US/09/900.237
; Current Filing Date: 2001-07-06
; Prior Application Number: 60/092,844
; Prior Filing Date: 18-07-14
; Prior Application Number: PCT/US99/15871
; Prior Filing Date: 1999-07-13
; Prior Application Number: 09/720383
; Prior Filing Date: 2000-12-21
Qy 641 --LLSKLUGGSRKNSAKKESDK--KKSCRHTDSTPVFNLDIEGVEGASFDE 693 ; NUMBER OF SEQ ID NOS: 33
Db 719 PCKWCICCCFGNKTCKTKTCKTSPKPEEKIKKKKENOAPAYALGIDEAPGA--ENE 776 ; SOFTWARE: Microsoft Office 97
Qy 694 KALIMSQSLERKGQSASVASTLMEENGVPSPATPENLKEAHVSCGYEDKSDWGM 753 ; SEQ ID NO: 4
Db 777 KASIVNQOKLEKKFGQSSVFPVASTLENGGTLSKASFASLKEAHVSCGYEDKOWGK 836 ; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Zea mays
; US-09-900-237-4
Query Match 69.1%; Score 3922; DB 10; Length 1148;
Best Local Similarity 67.2%; Pred. No: 0; Matches 734; Conservative 141; Mismatches 161; Index 56; Gaps 17;
Matches 734; Conservatve 141; Mismatches 161; Indexs 56; Gaps 17;
Db 4 EGEGTAGKOMKNTVPOTCOICSDMWGKTYDGRREVACDICSFVFCRCPCEYERKDQGSCP 63
Db 82 -GEPOYDKEVSHNHLPRLTSRQ--DTSGEFSAAASPERLVSSTSIAQGR--LPYSSDV 173
Db 64 OCTKTRYKLUKGSPAIKGDKDGLADECTIVEVNPQR--EISERALGWHITRGGEEM 120
Db 142 QCKTRKLUKGSPAIKGDKDGLADECTIVEVNPQR--EISERALGWHITRGGEEM 200
Db 201 DGVFQPFHPIP--NVLPLTINGQWDDIDPQDHALVP----SFVGGGKRIHPLPY-ADP 252
Qy 174 NQSPNRRRTVDP----VGLGNTAWKERYDVGKMKQKRNTGPVSTQASERGVYDAST 227
Db 253 NLVPQVPRSMPSKDLAAYGYGSVAKRKMESWKQKOR----MHTQTRNDGGDDDD-- 305
Qy 228 DILADEVALLNDEARQPLSRKVSISSRINPYRMIVMLVTCILHYRTNPVNFAF 287
Db 306 --ADLPLM-DEARQPLSRKIPLPSQINPYRMIVMLVTCILHYRTNPVNFAF 361
Db 288 WLWSVCIETWFAWSWILDQOPKWPVNREYDRLALRYDREGEPSOLAANDIFSTVDP 347
Db 362 WLWSVCIETWFAWSWILDQOPKWPVNREYDRLALRYDREGEPSOLAANDIFSTVDP 421
Qy 348 LKRPPLVTANTVLSIPLAVDYPVKVSCYVFDDGAAMLSFESLAETSEFARKWVPCCKYS 407
Db 422 LKEPPLVTANTVLSIPLSVDPVKVSCYVFDDGAAMLSFESLAETSEFARKWVPCCKYS 481
Qy 408 IERAPEWYFAAKIDYLKDQVTSFKDRRAMKREVEEKFTRINALVSKALKCPEGWVM 467
Db 482 LEPRAPENYFOOKIDYLKDQVTSFKDRRAMKREVEEKFTRINALVSKALKCPEGWVM 541
Qy 468 QDGTTPWPGNNTGDHPCMIVFLGONGQNLDAEGNEELPRLVVSREKRGFOHHKAGAMNA 527
Db 542 QDGTTPWPGNVRHGPQIVFLGQSGGHDVEGNEPLRVVVSREKRGFOHHKAGAMNA 601
Qy 528 LVRVSAVLTNAPYLNLDCDIYINNSKAIKEAMCFLMDPLGKVKCVOFPORFDGIDR 587
Db 602 LVRVSAVLTNAPYLNLDCDIYINNSKAIKEAMCFLMDPLGKVKCVOFPORFDGIDR 661
Qy 588 DRYANRNTVFEIDLNRGLDQGPGVYTGCFVNRYALGYEPPKVKHKEPS---- 640
Db 662 DRYANRNTVFEIDLNRGLDQGPGVYTGCFVNRYALGYEPPKVKHKEPS---- 718
Qy 641 --LLSKLUGGSRKNSAKKESDK--KKSCRHTDSTPVFNLDIEGVEGASFDE 693
Db 719 PCKWCICCCFGNKTCKTKTCKTSPKPEEKIKKKKENOAPAYALGIDEAPGA--ENE 776
Qy 694 KALIMSQSLERKGQSASVASTLMEENGVPSPATPENLKEAHVSCGYEDKSDWGM 753
Db 777 KASIVNQOKLEKKFGQSSVFPVASTLENGGTLSKASFASLKEAHVSCGYEDKOWGK 836
Qy 754 ECGWIGYGSYTEDILTGFEMHARGWRSTICMPKLAFKSAPINLSDRNLNOLVRLAWGSE 813
Db 837 DICWIGYGSYTEDILTGFEMHARGWRSTICMPKLAFKSAPINLSDRNLNOLVRLAWGSE 896
Qy 814 ILFSRHCIIWYGYNGRKLFLERPAVNTIYPTISIPLIMCPLAVCLFTNQFIPQIS 873
Db 897 IFFSNCHLWYGYGGGKFLERPSYINSIVPWTSIPLAVCTPLAICLTLGKFIPELN 956
Qy 874 NIASIWFSLFLSISIFATGILEMRWMSGVIDEMWRNEQFWIGGVSAHLFAVFOGIKVL 933

RESULT 6
US-09-900-237-10
Sequence 10, Application US/09900237
Patent No. US20020120124A1
GENERAL INFORMATION:
APPLICANT: Allen, Stephen
TITLE OF INVENTION: Plant Cellulose Synthases
FILE REFERENCE: BB1170 US CIP
CURRENT APPLICATION NUMBER: US/09/900, 237
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: 60/092, 844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US99/15871
PRIOR FILING DATE: 1999-07-13
PRIOR APPLICATION NUMBER: 09/720383
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Microsoft Office 97
SEQ ID NO 10
SEQ LENGTH: 1086
TYPE: PRT
ORGANISM: Zea mays
US-09-900-237-10

Query Match Similarity 68.3%; Score 3875.5; DB 10; Length 1086;
Matches 728; Conservative 145; Mismatches 161; Indels 57; Gaps 18;

QY 4 EGETACKPMKNTVPOTCOCISDNVKRDQFVACDICSPVCRCYCEYERKDGSNSCP 63
Db 23 DDPGPKPCKPREGONGQVQCICGDDVGLAGPGDPFVACNECAFPCRDCEYERREGIONCP 82

QY 64 OCKTRKRLKGSPAIGDKDDEGLADEGTVEENYP--QKEKITSERMLGHULTRK-GEEM 120
Db 83 QCKTRKRLKGQVHNLPRLSRQ--DTSGEFSAAASPERLSSVSTIAGKR--LPYSSDVN 174
QY 121 GEPQYKEVSHNLPRLSRQ--DTSGEFSAAASPERLSSVSTIAGKR--LPYSSDVN 174
Db 142 GAPQAO--LNPNVPLTNGQMDPPEQHALVPSFMG----GGKRKHPLPY-ADPS 193
QY 175 QSPNRRIUDP----VGLGNAWAKERDGHMKMOKRNTPVSTQASERGVQDIDASTD 228
Db 194 LPVQPRSMPSKDLAAYGGSVAKERMENWKQRER----MHOTGNDDGGDD-- 245

QY 229 ILADEALINDERAQPLSRKVSUPSSINPYRMVIMRLVTCLEHYRITNPVNAFLW 288
Db 246 --ADPLM--DEARQQLSKRKLPLSSQINPYRMVILILVVLGFPHYRVMHPVNDALW 302

QY 289 LVSVICEIWALSWILDQFPKWPVNPRTYLDLRLRYDRIGESEPSOLAVIDFVSDP 348

Db 303 LJSVICBIIWFAMSWLQDFPKWPPIERETYDRLSLRFKQGQPSOLAPIDFFSVNDPL 362

QY 349 KEPPLVNTVMSLAVDYPVDKVSCVFDGAMLSFESLAETSFARKWPFCKKSY 408
Db 363 KEPPLVNTVMSLAVDYPVDKVSCVFDGAMLSFESLAETSFARKWPFCKKSY 422

QY 409 EPRAPENYFAAKIDYLKDVKVOTSEFVDRRAMREYEEFKIRINALVSKALKCPEEWWMQ 468

RESULT 7
US-09-900-237-8
Sequence 8, Application US/09900237
Patent No. US20020120124A1
GENERAL INFORMATION:
APPLICANT: Allen, Stephen
TITLE OF INVENTION: Plant Cellulose Synthases
FILE REFERENCE: BB1170 US CIP
CURRENT APPLICATION NUMBER: US/09/900, 237
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: 60/092, 844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US99/15871
PRIOR FILING DATE: 1999-07-13
PRIOR APPLICATION NUMBER: 09/720383
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Microsoft Office 97

SEQ ID NO 8
LENGTH: 1165
TYPE: PRT
ORGANISM: Zea mays
US-09-900-237-8

Query Match Similarity 67.7%; Score 3843.5; DB 10; Length 1165;

Db 423 EPRAPENYFQQKIDYLKDVKVAFRRAMREYEEFKIRINALVAKQVPEEGTMQ 482
QY 469 DCTPWPGNNTGDHPGMQVFLQGQNGGLDAEGLNPLRVYVSREKRCGQHHKAGMNL 528
Db 483 DCTPWBCNNVRHPGMQVFLQGQGLDCENEPRYREKRCGYNHKKAGMNL 542
QY 529 VRVSAVLTNGPFTLNLDCHYTNNSKALREAMCFMDPNLGKQVCYQFPOFDDGTDKND 588
Db 543 VRVSAVLTNAPTLINLDCDHYTNNSKAIKEAMCFMDPLGKVCYQFPOFDDGTDKND 602
QY 589 RYANRNYFEDINLRGLDGIOPVYGTGCVENRATLYYRPIKHKKSSLKLUG- 647
Db 603 RYANRNYFEDINMKGLDGIOPVYGTGCVFRRQALYDPR--KTKPP--SRWCNC 657

QY 648 -----GSRKKSKAKKEDKKK--SQRHTDSTPVFNLDIEEGEGFDEK 694
Db 658 WPWKWFCGCCCGCFCGKTTKKEKKLKFPLKKEPAKPAKGSAPLNLSDRUNQVLRWALGVEI 715

QY 695 ALIMMSQMSLEKREGQSAVFVASTLMEENGVPSSATPENLKAHVITSCGYEDDSWGMG 754
Db 716 AGIVNQOKLEKKFGQSSVFVSTILLENGTLKSASPLKRAIHVISCGEDKTDMGK 775

QY 755 IGWIYGSWTEDILTGETNMHARGWRSITCMPKPAFKGSAPILNLSDRUNQVLRWALGVEI 814
Db 776 IGWIYGSWTEDILTGFKMHCWRSIVCIPKRAFKGSAPILNLSDRUNQVLRWALGVEI 835

QY 815 LESRHCPIWIGNGRLKFLERAYVNNTIPISTSIPUMCYCLAVLFTNQFIQISN 874
Db 836 IFSNHCPIWYGGGLFLERASYINSIVYPTISIPLACLTPLACLLCRKTFITBLN 895

QY 875 IASIWFLSLSIFATGITLEMRSVGIDEMWNEQFWVIGVSAHLFAVFOGLKVLG 934
Db 896 VASLWFMSLFICIFATSILEMRWSVGIDDMWRNEQFWVIGVSAHLFAVFOGLKVLG 955

QY 935 IDTNFTVTSKASDEGDFAELTLFKWMTLIPPTTLIIVNNGVWAGVSAINSQSW 994
Db 956 VDSTFTVSKGDD--DEEFSELYTFKWTTLIIPPTTLIIVNNGVWAGVSAINSQSW 1014

QY 995 PLEGKLFAPWVHLPKLGKLMGRNRPITVWVSVLASSIFSLWVDRDPFSTRV 1054
Db 1015 PLEGKLFAPWVHLPKLGKLMGRNRPITVWVSVLASSIFSLWVDRDPFSTRV 1074

QY 1055 GDPILEGGINC 1065
Db 1075 GPLLECGGLDC 1085

Best Local Similarity 66.3%; Pred. No: 0; Matches 727; Conservative 134; Mismatches 178; Indels 57; Gaps 17;

Qy 2 ESEGETACKPKMKNIVPQTQICSDNVGKTYDGRFVACDICSPVCPYCYERKDGMQS 61
Db 94 DRESGAAGGGAARRAEADPCQICDEVGVGFDFGRFVAECAPVCRAYEYERREESQQA 153

Qy 62 CPOCKTRYKLKGSPAIQDKDEGLAD-----EGTVEFNPQKEKISERMLGWHLT 113
Db 154 CPOCRTTRYKLKGCPRVGDEEDGVDDLEGERGLQDAHEIDPQ--YVAEMLRAGMS 211

Qy 114 RGKGEEMBSPQYKKEVSHNLPRLTSR---DTSGEFSASPERLSYSTIAGKR--L 167
Db 212 YGRGGD-AHGFSPV-- PNVPLTINGQMDPQHAEPEQHAIWPSYMSGG--GGGKRHPL 265

Qy 168 PYSSDVNQSPNRRIVDP--- VGLGNWAKWKRVDGMKMKOKNTGVSSTOASERGV 221
Db 266 PF-ADPNLFVQPRSMDPKDLAAYGYSWAKWERMEGRKKQKR---LQIVRSEGG 318

Qy 222 DIDASTDILADEALLNDARQPLRSRKVSISSRINPYRMVIRLVLICLFLHYRITNPV 281
Db 319 DWGDDD--ADPLM-DEARPLSRKVPISSRINPYRMVIRLVLGFFEHYRMPA 374

Qy 282 PNALFALWIVSVCIEIWFAWSWILDQFPKWPVNRTEYLRLALRYDRGEPSOLAVIDF 341
Db 375 KDAFLWISVICIEWFAMSWILDQFPKWLPIERETYLRLSLRFKGQPSOLAPIDF 434

Qy 342 VSTVDPKKEPPVANTVLSILAVDYPDKVSYVFDDGAMLSFESLAETSFARKWP 401
Db 435 VSTVDPKKEPPVANTVLSILSVDPKVEKSYVSDGGAATIFEALESFARKWP 494

Qy 402 FCKKYSTEPPAREPWYAAKIDYLKDQVTSFYKDRRAMKREYEEFKRINALYSKALKCP 461
Db 495 FSKKFNTIERPAPENWFOOK-IDYLKDVKVASFVRERRAKMREYEEFKRINALVAKWP 554

Qy 462 EGGVMQDGTPWIGNNIQGDHPGMQIVFGQNGGLDAEGNELPRLVYSSREKPGFQHKK 521
Db 555 EBGWTMOPGSWPWGNVNRDHPGMQIVFGQNGGLDAEGNELPRLVYSSREKPGYNHKK 614

Qy 522 AGAMNALYRVSAVLNTGPFITLNIDCDHYTINNSKALREAMCFMLMDPNLGQCVWOFPOF 581
Db 615 AGAMNALYRVSAVLNSAAYLNIDCDHYTINNSKALREAMCFMLMDPLVKGKCVQOFPOF 674

Qy 582 DGIDKNDRYANRNFEDINLRGLDGTOPGVVYGTGCYFNRTALGYEPPIKVKKHPS- 640
Db 675 DGIDKNDRYANRNFEDINMKGDLDGTOPGVYGTGCYFRRQALYGDAP---KTKKPPS 731

Qy 641 -----LLSKC - GGSRKNSKAKKESDKKKSGRHTDSVPUFUDDEEGVCG 689
Db 732 RTNCNCWPWKMCSCSRNKKKTPKTEKKRFLFKAENSPAVALEGEDEGAGA- 790

Qy 690 FDEDEKALIMQSLEKRTGQSAVFVASTLMEENGVPPSATPENLKEAHVTCGYEDS 749
Db 791 -DIEKAGIVNQKLEKKEQQSYFVASTLLENGTLKASAPASLKEAHVSCGYEDK 849

Qy 750 DWMEIGMIGYGSVETDILGKHMARGWSRIVCMPKLPKAFKGASPAPINSLDRNQVWLW 809
Db 850 DWKGEIGWIGYGSTEDILGKHMCHGWRISYICPKRPAFKGASPAPLNLSDRLHQVRLWAL 909

Qy 810 GSVEILSRHCPWYGINGRKLERFVYNTIYTPTISIPLUMCULLAVLFTNOFI 869
Db 910 GSVEIFFSKHCPWYGGGLKFLERFSYIIVSPWHSIPLAYCTPLPATCOLTGKPP 969

Qy 930 KVLAGIDNTFTWIKSKASDDEGDAEFLYLFKWTLLIPPTLUVLNUVWAGVSYAING 989
Db 1030 KVFAIGDTSFTWIKSKAD-DEEFSELTYFWKWTLLIPPTLULINFVAGISNATING 1088

Qy 990 YOSWGPFLGKLFAFWVWHLWPLKGJLMGRQNRTPTIVWWSVLLASIFSLWWRDPF 1049

RESULT 8
US-09-900-237-14
; Sequence 14, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIORITY NUMBER: 60/092,844
; PRIORITY NUMBER: 60/092,844
; PRIORITY FILING DATE: 1998-07-14
; PRIORITY APPLICATION NUMBER: PCT/US99/15871
; PRIORITY FILING DATE: 1999-07-13
; PRIORITY APPLICATION NUMBER: 09/720383
; PRIORITY FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 1039
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (201)
; US-09-900-237-14

Query Match Best Local Similarity 67.0%; Score 3804.5; DB 10; Length 1039;
Matches 725; Conservative 117; Mismatches 160; Indels 75; Gaps 16;

Qy 5 GETAKPKMKNIVPQTQICSDNVGKTYDGRFVACDICSPVCPYCYERKDGMNQSCPO 64
Db 22 GHEPPKALKNLDQVCEICGDKGWDLQDFLYVQFCNECGFPVCPYCYERREGHSICPQ 81

Qy 65 CKTRYKRKGSPLAIGDKDEGLADEGTVEFNPQKE---K1SERMGWHLTRKGKEM 120
Db 82 CKTRYKRKGSPLAIGDKDEGLADEGTVEFNPQKE---K1SERMGWHLTRKGKEM 120

Qy 121 GEPOQDKEVSHNHLPLRITS -- RDTSGFSAASPERLSVSSTIAGKRLPVSDDVNGSPN 178
Db 140 -----DDDNQEPPTVLAGRSRSPVSGF-----PISSNAYGDML-SSSLKRVH 184

Qy 179 RRYVDPVGLN--- VAWKERYDWGMKQEKNTGPVSTOASERGVVDIDASTDILAD 232
Db 185 PYPVSEPCSSARWDEKKXDGWKDMDWKL-QDQNGL-----EPPSDP 227

Qy 233 EALINDEARQPLRSRVSFSSRINPYRIMVRLVCLGLFLAYRITNEVNAFLAWVSV 292
Db 228 AAML-DEARPLSRKVNPASSKINPYRIMVRLVCLGLFLAYRITNEVNAFLAWVSV 286

Qy 293 ICEIWFASWILDQFPKWPVNRTEYLRLALRYDRGEPSOLAVIDFVSTVDPKLEPP 352
Db 287 ICEIWFASWILDQFPKWPVNRTEYLRLALRYDRGEPSOLAVIDFVSTVDPKLEPP 346

Qy 353 LVTANTVSLAVDYPDKVSYVFDDGAMLSFESLAETSFARKWPCKYSEPPRA 412
Db 347 LYVANTVSLANDYDYPDKVSYVFDDGAMLSFESLAETSFARKWPCKYSEPPRA 406

Qy 413 PEVYFAKIDYLKDVKVSYVFDDGAMLSFESLAETSFARKWPCKYSEPPRA 472
Db 407 PEVYFAKIDYLKDVKVSYVFDDGAMLSFESLAETSFARKWPCKYSEPPRA 466

Qy 473 WPCNNTGHGPGMQLVFGQNGGLDAEENELPRLVYSSREKPGFQHKKAGAMNALVRV 532
Db 467 WPCNNTGDHPGMQLVFGQNGGLDTEGNOLPRLVYSSREKPGFQHKKAGAMNALVRV 526

QY 533 AVLNGPFLNUCDWYINNSKALRGRACMFLMDPNIGKQCVYQFFQRFDGIDKNDRYAN 592
Db 527 AVLNTNPFPMLNCLMDCHYVNNSKARGRACMFLDPQGKVKYQFQRFDGIDKDRYAN 586
Qy 593 RNTVFFDINLRGLDGIOGPVYGTGCVENRFTALYGEPPIKVKKRPSLLS--KLCGS 649
Db 650 RKKNSRAKKEEDKKSGRHTSTVPUFNLDIEEGTEGAGFDEKALLMSOMSLERFGQ 709
Qy 645 RKYY-----KEKIN-----DANGEAASLKGMDDKLTSQNFKEKFGQ 683
Db 710 SAVVFASLTMEENGYPAPSATPENLKEAIIHVISCGYEDKSDMGMEIHWIGSYTEDILTG 769
Db 684 SSIFVSTMLMEEGGVPPSSPAALKEAIHVSCGYEDKTBGLEGWIQSITEDILTG 743
Qy 770 FKMHAIGWRSITYCMPKPLPAFKGSAPNLSDRLNQVRAWSVEILFSRHCPIWYSG-NG 828
Db 744 FKMHCIGWRSITYCMPKPLPAFKGTAPTNLSDRNQVRAWSIEFFSHCPIWYSGFKEK 803
Qy 829 RLKFLLRFAYANTTVPFTSIPLVAVCILPAVCLLTDKFIMPISPTAGLYFVALPSSII 863
Db 804 KLKWLWERFAYANTTVPFTSIPLVAVCILPAVCLLTDKFIMPISPTAGLYFVALPSSII 863
Qy 889 ATGILEMRWSSGVGIDEMEARNEOFWVGGVS AHLFAVFOGLIKVLAGSDTNTFTVSGASDE 948
Db 864 ATGILELKWSVCSVSEIWNRNSEQFWVGGVS AHLFAVFOGLIKVLAGSDTNTFTVSGASDE 922
Qy 949 DGDGFATEYLFRKWTLLIPPTTLIVLNUVWVAGVSTAINGYOSWGPLGKLFKLMGRQNT 1008
Db 923 DEEFGGLYTFKWTLLIPPTTILINIVGVWVAGSAINNGYOSWGPLGKLFKLMGRQNT 982
Qy 1009 HLYPFKLGMGRQNTPTIVVWVSVLIALSISFLSLWWRIDPFTSRVGPDTILECGINC 1065
Db 983 HLYPFKLGMGRQNTPTIVVWVSVLIALSISFLSLWWRIDPFTSRVGPDTILECGINC 1039

RESULT 9
US-03-900-237-32
; Sequence 32, Application US/09900237
; Patient No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: B1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900, 237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092, 844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 32
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
; US-03-900-237-32

RESULT 9
US-03-900-237-32
; Sequence 32, Application US/09900237
; Patient No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: B1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900, 237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092, 844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 32
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
; US-03-900-237-32

Query Match 61.4%; Score 3483.5; DB 10; Length 701;
Best Local Similarity 90.9%; Pred. No. 3.5e-291;
Matches 637; Conservative 42; Mismatches 21; Indels 1; Gaps 1;

Qy 366 DYPVDKVSYYFDGAMLSPESLAPMSEFARKWVFPCKKSYIPEPAPEWFAAKDYLK 425
Db 1 DYPVERKVSKYSDGAMLTFEALSETSEFARKWVFPCKKSYIPEPAPEWFAQKIDYLK 60
Qy 426 DKVQTFPKVDRAMKKEYEKEKIRIHALVSLALKCREGWYMQDCPWPWGNNTGDPGMI 485
Db 61 DKVQTFPKVDRAMKKEYEKEKIRIHALVSLALKCREGWYMQDCPWPWGNNTGDPGMI 120

QY 486 QVILGONGGLDAEKGELPRLVY'SREKRGFQHKKAGAMNALVRSAVLTGFLNL 545
Db 121 QVFLQSGSLDAEKGELPRLVY'SREKRGFQHKKAGAMNALVRSAVLTGFLNL 180
Qy 546 CDHYINNSKALRENCMFLDNPNGKQCVYQFOPRFDGIDKDRYANRNTVFFDINRL 605
Db 181 CDHYINNSKALRENCMFLDNPNGKQCVYQFOPRFDGIDKDRYANRNTVFFDINRL 240
Qy 606 DGTOPPVWVGTGCVENRFTALYGEPPIKVKKRPSLLS--KLCGS 664
Db 241 DGTOPPVWVGTGCVENRFTALYGEPPIKVKKRPSLLS--KLCGS 300
Qy 665 SGRITDSTPVFVNLDDEBVGEGFDDERKALMSOMSLERFGQSAVVFASLTMEENGV 724
Db 301 SGKHVDSTPVFNFNLDIEBVGEGFDEKSLLMSOMSLERFGQSAVVFASLTMEENGV 360
Qy 725 PSATPENLKEAIIHVISCGYEDKSDMGMEIHWIGSYTEDILTGKFMHARGHRSIYCP 784
Db 361 PQSATPETLUKEAIIHVISCGYEDKSDMGMEIHWIGSYTEDILTGKFMHARGHRSIYCP 420
Qy 785 KLPFKGSAPNLSDRLNQVRLWALGSVEILFSRHCPIWYSGRKLWERAVNTTIV 844
Db 481 PTAIPPLMCTIPAVCILTNKTFIQSISNLASIWFISLSPSFATGFLKMKRNGVQD 540
Qy 905 WWRNEQFWVIGYSAHLFAVFOGLIKVLAGSDTNTFTVSKASBDGDFAELEYLFWKTLL 964
Db 541 WWRNEQFWVIGYSAHLFAVFOGLIKVLAGSDTNTFTVSKASBDGDFAELEYLFWKTLL 600
Qy 965 IPPPTLIVNLUVWVAGVSYAINGYOSWGPLGKLFRAFWVWVHLYPFLKGMRQNT 1024
Db 601 IPPPTLIVNLUVWVAGVSYAINGYOSWGPLGKLFRAFWVWVHLYPFLKGMRQNT 660
Qy 1025 PTIVVWWSULLASISFLSLWWRIDPFTSRVGPDTILECGINC 1065
Db 661 PTIVVWWSULLASISFLSLWWRIDPFTSRVGPDTILECGINC 701

RESULT 10
US-09-838-539-6
; Sequence 6, Application US/09838539
; Patient No. US20020129401A1
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/03/05
; CURRENT APPLICATION NUMBER: US/09/838, 539
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029, 987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960, 048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 6
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Gossypim hirsutum
; US-09-838-539-6

Query Match 61.2%; Score 3473.5; DB 10; Length 974;
Best Local Similarity 62.1%; Pred. No. 4.3e-290;
Matches 660; Conservative 129; Mismatches 176; Indels 97; Gaps 15;

Qy 12 MKNIVPQCOICSDNVKGKVDGDRFVACDCSFVCRCPYERKRDGNOSCPCKTRYR 71
Db 2 MESGVP-VCHTCGBHVGLVNLNGPFPVACHECNFPICKSCFEPYDKEGRKACRC----- 54

RESULT 14
US-09-838-539-7
; Sequence 7, Application US/09838539
; Patent No. US20020159401A1
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15/21/03/US
; CURRENT APPLICATION NUMBER: US/09/838, 539
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029, 987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960, 048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Goosypium hirsutum
; US-09-838-539-7

Query Match 49.4%; Score 2803; DB 10; Length 685;
Best Local Similarity 74.5%; Pred. No. 1.5e-232; Mismatches 512; Conservative 77; Indels 22; Gaps 6;

Qy 396 ARKWPFCKKYSEPRAPWFAKIDYLKDQTSFWKDRRAMEKREYEFKRINALWS 455
Db 1 ARWWPFCKKHNPRAPEFVYNEKIDYLKDQYKERRAMREYEFKRINALWA 60

Qy 456 KALKCPERGWMDGTPWPGNNNGDHPGMIOVFLQGONGGLDAGNEELPRLYVSREKRP 515
Db 61 KAKQKPERGWMDGTPWPGNNNGDHPGMIOVFLQGSGGLDIDNEPLRUYVSREKRP 120

Qy 516 FOHHKKAGMNALVRSAVLTINGPFILNDCDHYINNSKALREAMCFLMDPNLGKQCY 575
Db 121 YOHHKKAGMNALVRSAVLTINGPFILNDCDHYINNSKALREAMCFLMDPNLGKQCY 180

Qy 576 QPPRGDKNDRYANNTVFDINLRLDGQIOPWYGTGGCFNRTALYYPEPKV 635
Db 181 QPPRGDGDIDRHDYANRNVFDINMLGDLQGPVYVGTCVFNRLGYDPYSEK 240

Qy 636 HKK-----PSLISKCGSRKNSKAKE-----SDKKSGRH--TDSTVPFN 677
Db 241 RPKMTDCWPSSWCCCCCGSRKNSKKKGKEKGKGLGGLIYGGKKKMGKNNVKVKGSAVFD 300

Qy 678 LDIEEGVGAGEDD-EKALLMSOMSLERKFGOSAVESTLMEENGGVPPSATPENLKE 736
Db 301 LEEFEGLE-GYEELERSTLMSOKNFERKFGOSPVFTASTLMEENGGLPEGTNSTSLIKE 358

Qy 737 ATIVISCYEDKDWGMIGWIYGSVIEDILGFKMHARGWRSIYCMPKLPFKGSAVIN 796

RESULT 15
US-09-900-237-16
; Sequence 16, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900, 237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092, 844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Glycine max
; US-09-900-237-16

Query Match 46.6%; Score 2647.5; DB 10; Length 610;
Best Local Similarity 78.8%; Pred. No. 3.1e-219; Mismatches 482; Conservative 55; Indels 3; Gaps 2;

Qy 455 SKALKCPERGWMDGTPWPGNNNGDHPGMIOVFLQGONGGLDAGNEELPRLYVSREKRP 514
Db 61 AKAQKPERGWMDGTPWPGNNNGDHPGMIOVFLQGSGGLDIDNEPLRUYVSREKRP 60

Qy 515 GFOHHKKAGMNALVRSAVLTINGPFILNDCDHYINNSKALREAMCFLMDPNLGKQCY 574
Db 61 GFOHHKKAGMNALVRSAVLTINGPFILNDCDHYINNSKALREAMCFLMDPNLGKQCY 120

Qy 575 VOPQRFGDKNDRYANNTVFDINLRLDGQIOPWYGTGGCFNRTALYYPEPKV 634
Db 121 VOPQRFGDGDIDRHDYANRNVFDINMLGDLQGPVYVGTCVFNRLGYDPYSEK 180

Qy 635 KHKPLSLISKCGSRKNSKA-KKEDDKKSKSRHHTSTVPUFNLDDETEEGVAGA 693
Db 181 EDLEPNIVKSCCGSRKKGKGKNNKSYKKKAMGRTESTVPUFNMEIEEGE-GYDDE 238

Qy 694 KALIMSONSLEKRGFGOSAVESTLMEENGGVPPSATPENLKEAHVISCYEDKSDWMG 753
Db 239 RTLLMSOKSLEKRGFGOSPVFTASTLMEENGGLPEGTNSTSLIKE 298

Qy 754 EIGWIGSYTEDILGFKMHARGWRSIYCMPKLPFKGSAVIN 813

Db	299	EIGWIGSVTEDILTGKMHARGWISIYCMPPRAFKGASPINALSDRLNQVLRWALGSIE	358
Qy	814	IIFSRICPIIWGYNGRKFLERPAYVNITIPTISPLMCLLAVLFTNFIQIS	873
Db	359	IFLSRICPILWGYNGRKFLPLMLRAYINIVFPTSIPLIAYCPLPAFCULTNKFIPES	418
Qy	874	NIASIMFLSLSLTSIAFGILEMRKWGVGIDEMARNEQFWVIGVS AHLFAVGOFILKYLIA	933
Db	419	NFASMWMILLFSISIFTSILERLRWSGVSIEDMWNRNEQFWVIGGTSAHLFAVGOFILKYLIA	478
Qy	934	GIDTNFVUTSKASDEDDGDAEFLYKFMWLLJPPTELLIVNLYGVAGVSAYINSQSW	993
Db	479	GIDTNFVUTSKASDEDDGDAEFLYKFMWLLJPPTELLIVNLYGVAGVSAYINSQSW	538
Qy	994	GPLFGKLFFAFWVIVHLYPFLKGLMGRNRTPTIVVVWSVLLASIFSLWLRIDPFTSRV	1053
Db	539	GPLFGKLFFAFWVIAHLYPFLKGLLGRNRTPIIVVWSVLLASIFSLWLRIDPFTSDS	598
Qy	1054	TGPDILCGINC 1055	
Db	599	NKLINQGCGINC 610	

Search completed: June 16, 2003, 11:15:47
Job time : 36 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2003, 11:05:53 ; Search time 18 Seconds

Perfect score: 5677 ; (without alignments)
Sequence: 1 MESEGETAGPKPMKNNIVPQTC.....IDPFITSRVTGPDPDILECGINC 1065

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422322 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

- 1: /cgn2_6/pctdata/1/1aa/5A_COMB_pep:*
- 2: /cgn2_6/pctdata/1/1aa/5B_COMB_pep:*
- 3: /cgn2_6/pctdata/1/1aa/6A_COMB_pep:*
- 4: /cgn2_6/pctdata/1/1aa/6B_COMB_pep:*
- 5: /cgn2_6/pctdata/1/1aa/PCTRUS_COMB_pep:*
- 6: /cgn2_6/pctdata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4133.5	72.8	881	4 US-08-960-048-8	Sequence 8, Appli
2	3473.5	61.2	974	4 US-08-960-048-6	Sequence 6, Appli
3	2803	49.4	685	4 US-08-960-048-7	Sequence 7, Appli
4	278	4.9	693	4 US-08-960-048-11	Sequence 11, Appli
5	245.5	4.3	861	4 US-08-960-048-12	Sequence 12, Appli
6	236.5	4.2	756	4 US-08-960-048-10	Sequence 10, Appli
7	236.5	4.2	756	4 US-09-147-236-3	Sequence 3, Appli
8	28.5	4.0	3031	1 US-07-689-008-3	Sequence 2, Appli
9	196	3.5	723	4 US-08-960-048-9	Sequence 9, Appli
10	196	3.5	723	5 PCT-US1-01726-4	Db
11	124	2.2	357	1 US-08-119-773-4	Sequence 4, Appli
12	116	2.0	357	1 US-08-119-773-6	Sequence 5, Appli
13	115	2.0	346	1 US-08-119-773-5	Sequence 3, Appli
14	115	2.0	357	1 US-08-119-773-2	Sequence 2, Appli
15	110	1.9	1587	4 US-09-000-094-46	Sequence 46, Appli
16	107	1.9	686	2 US-08-768-301-4	Sequence 4, Appli
17	106.5	1.9	1437	3 US-09-061-400-2	Sequence 2, Appli
18	106.5	1.9	1453	2 US-09-001-273-2	Sequence 2, Appli
19	106.5	1.9	1453	4 US-08-843-459A-2	Sequence 2, Appli
20	104	1.8	1167	1 US-08-485-568A-6	Sequence 2, Appli
21	104	1.8	1167	2 US-08-590-554A-6	Sequence 6, Appli
22	104	1.8	1167	2 US-09-184-223-6	Sequence 6, Appli
23	104	1.8	3033	1 US-07-925-635-5	Sequence 5, Appli
24	103	1.8	1168	1 US-08-120-717A-9	Sequence 9, Appli
25	102.5	1.8	1871	2 US-08-594-869-1	Sequence 1, Appli
26	102.5	1.8	1871	3 US-09-449-546-1	Sequence 1, Appli
27	101.5	1.8	765	2 US-08-825-886-19	Sequence 19, Appli
28	101.5	1.8	900	1 US-08-425-061-20	Sequence 20, Appli
29	101.5	1.8	900	2 US-08-425-886-20	Sequence 21, Appli
30	101.5	1.8	914	1 US-08-425-061-21	Sequence 22, Appli
31	101.5	1.8	914	2 US-08-825-886-21	Sequence 23, Appli
32	101.5	1.8	914	2 US-08-825-886-22	Sequence 24, Appli
33	101.5	1.8	1202	1 US-08-825-886-22	Sequence 2, Appli
34	101.5	1.8	1202	2 US-08-825-886-22	Sequence 2, Appli
35	101.5	1.8	1363	1 US-08-825-886-23	Sequence 2, Appli
36	101.5	1.8	1363	2 US-08-825-886-23	Sequence 2, Appli
37	101.5	1.8	1852	1 US-08-425-061-24	Sequence 2, Appli
38	101.5	1.8	1852	2 US-08-825-886-24	Sequence 2, Appli
39	101.5	1.8	1863	1 US-08-598-591-2	Sequence 2, Appli
40	101.5	1.8	1863	1 US-08-480-784-2	Sequence 2, Appli
41	101.5	1.8	1863	1 US-08-480-553-2	Sequence 2, Appli
42	101.5	1.8	1863	1 US-08-487-002-2	Sequence 2, Appli
43	101.5	1.8	1863	1 US-08-483-556B-2	Sequence 2, Appli
44	101.5	1.8	1863	1 US-08-798-691-2	Sequence 2, Appli
45	101.5	1.8	1863	1 US-08-798-691-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-960-048-8

; Sequence 8, Application US/08960048C

; Patent No. 6211443

; GENERAL INFORMATION:

; APPLICANT: Stalker, D. et al.

; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter

; TITLE OF INVENTION: Sequences

; FILE REFERENCE: 15621/01/US

; CURRENT APPLICATION NUMBER: US/08/960, 048C

; CURRENT FILING DATE: 1997-10-29

; PRIORITY APPLICATION NUMBER: 60/029, 987

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 8

; LENGTH: 881

; TYPE: PRT
; ORGANISM: Oryza sativa

; US-08-960-048-8

; Query Match

; Best Local Similarity

; Pred. No: 0;

; Matches

; Conservative

; Mismatches

; Indels

; Gaps

; 3;

; Description

; Sequence 8, Appli

; Sequence 6, Appli

; Sequence 7, Appli

; Sequence 11, Appli

; Sequence 12, Appli

; Sequence 10, Appli

; Sequence 3, Appli

; Sequence 2, Appli

; Sequence 9, Appli

; Sequence 4, Appli

; Sequence 7, Appli

; Sequence 5, Appli

; Sequence 3, Appli

; Query Match

; Best Local Similarity

; Pred. No: 0;

; Matches

; Conservative

; Mismatches

; Indels

; Gaps

; 3;

; Description

; Sequence 8, Appli

; Sequence 6, Appli

; Sequence 7, Appli

; Sequence 11, Appli

; Sequence 12, Appli

; Sequence 10, Appli

; Sequence 3, Appli

; Sequence 2, Appli

; Sequence 9, Appli

; Sequence 4, Appli

; Sequence 7, Appli

; Sequence 5, Appli

; Sequence 3, Appli

; Sequence 2, Appli

; Sequence 46, Appli

; Sequence 4, Appli

; Sequence 2, Appli

; Sequence 1, Appli

; Qy

</div

Db 361 LDCDHYINNSKALRACMCFMLDPNIGRSVCYVQFFQRDFGIDRNDRYANRNTVFEDINLR 420
 Qy 604 GLDGQGPVYVGTVGTCVNFRNTALGYEPPIKVKHKPKPSLISKLGCSRKRKSKAKESDK 663
 Db 421 GLDGQGPVYVGTVGTCVNFRNTALGYEPPIKVKHKPKPSLISKLGCSRKRKSKAKESDK 479
 Qy 664 KSGRHTDSTPVFNIDDDIECVERGAGFDDKEKKALMSQMSLERKFGOSAVFVASTLMENG 723
 Db 480 KSNKHDASVNPENNEDEIEGGVEAGGFDEKLSLMSQMSLEKRFQSAFVASTLMEGG 539
 Qy 724 VPPSAPENILKEAHTHVISCGYEDKSDWMBIGWYGVTHDILTCFKMHARGWRSIYCM 783
 Db 540 VPOSTTPESLKEAHTHVISCGYEDTREWGEIGWLYGSVTEIDLTCFKMHARGWRSIYCM 599
 Db 784 PKLPFKGSAPINSLSDLRNLQVRLWALGSVETLFSRHCPIWYGYGRKFLERFAYNTTI 843
 Db 600 PKRPAFKGSAPINSLSDLRNLQVRLWALGSVETLFSRHCPIWYGYGRKFLERFAYNTTI 659
 Qy 844 YPTISPLIMCTLLAVCLFTNOFTIPOISNIASWFLSFLSIFTANGLEMRSVCGID 903
 Db 660 YPLTSPLIYCVBACILTGKFPIESENFASTWFLSIFTANGLEMRSVCGID 719
 Qy 904 EWWREOFWVIGGYSAHLRAFVFOGTLKVLAGIDTNTWKAASDEGDFDALLYLKWTLL 963
 Db 720 EWWREOFWVIGGYSAHLRAFVFOGTLKVLAGIDTNTWKAASDEGDFDALLYLKWTLL 779
 Qy 964 LIPTTLLIVLNWVGWAGSYAINGYQSMPLGKLFFAEWVWVLYPELKGLMQRNR 1023
 Db 780 LIPTTLLIVLNWVGWAGSYAINGYQSMPLGKLFFAEWVWVLYPELKGLMQRNR 839
 Qy 1024 TPTIVVVWSULLASFLSLWVRRIDFTSRVGPDTLECGINC 1065
 Db 840 TPTIVVVWSULLASFLSLWVRRIDFTSRVGPDTLECGINC 881

RESULT 2
 US-08-960-048-6
 ; Sequence 6, Application US/08960048C
 ; General Information:
 ; Applicant: Stalker, D. et al.
 ; Title of Invention: Plant Cellulose Synthase and Promoter
 ; Title of Invention: Sequences
 ; File Reference: 15621/01/US
 ; Current Application Number: US/08/960,048C
 ; Current Filing Date: 1997-10-29
 ; Prior Application Number: 60/029,987
 ; Prior Filing Date: 1996-10-29
 ; Number of SEQ ID Nos: 12
 ; Software: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 974
 ; TYPE: PRT
 ; ORGANISM: Gossypium hirsutum
 ; US-08-960-048-6

Query Match 61.2%; Score 3473.5; DB 4; Length 974;
 Best Local Similarity 62.1%; Pred. No. 0;
 Matches 660; Conservative 129; Mismatches 176; Indels 97; Gaps 15;

Db 12 MKNTVPOTCOICSDNVGKTVDGDRREVACICSFPCVCPYCEYERIDGNOCPOCHTRYKR 71
 Qy 1 : | : | - | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 2 MESGPV-VCHTCGERVGLNVNGEPFVACHECNFPICKSCFEYDLREGKACLRC----- 54
 Db 72 LKGSPAIPIGDKEGLADECTVERNPYQKERKISEMLGLWHLTRGKGEEMGPQYKEVSH 131
 Qy 1 : | : | - | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 55 --GSP---YDENLDD-----VERATGDOSTMALH--NISQDVG-----IHA 90
 Qy 132 NHLPDRTSRQDTGFSAAASPERLVSSTIAGKRLPYSSDVNQSPNRRIVDPVGLGNA 191
 1 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 91 RHSSVSTLSEMAEDN-----GNSI 111

RESULT 3
 US-08-960-048-7
 ; Sequence 7, Application US/08960048C
 ; General Information:
 ; Applicant: Stalker, D. et al.
 ; Title of Invention: Plant Cellulose Synthase and Promoter
 ; Title of Invention: Sequences
 ; File Reference: 15621/01/US
 ; Current Application Number: US/08/960, 048C
 ; Current Filing Date: 1997-10-29

Qy 192 WKRDVGKMKQEKNTGPySTOASeRGWVDASTDILADEALND-----EAROPLSR
 112 WKNRVESEWEKEKKNNKKKPATKVERAE-----IPPEOQMEDKIPADPASQPLST 161
 Db 248 VSIPSSRINPRYMIWMLRVLWLCFLHLRINTNPWNAFLWLWVSEWFWALWSWLD 307
 Qy 308 PKNPVNPNTYTYLALRDRGEPSPOLAADVDFSTVDPULKEPLVLTANTVSLAVD 367
 Db 222 PKWVNPNTYDLSARVEREPEGDELAADFVSTVDPFLKEPLITANTVSLALDY 281
 Qy 368 PVKUVCYEDDGAISSPESLATEFARKWVFCCKSYIERRAPEWFAAKTDYLK 427
 Db 282 PVKVSCYISDDGGRAMLTFSVLTADFAKWKWPCKFSIEPRAPEFYSOKDYLK 341
 Qy 428 VOTSFVKDRAKMRYEERFIRNALVSKALKCPEEGWWMDGTPWPGNNTGDHGMQV 487
 Db 342 VQSFVKERRAMRDYEERKIRNALVAKQKTDGEWMQDSWPGNPRDHGMQV 401
 Qy 488 FLGONGGLDAGEGNELPRLVYVSREKPFQOHKKAGAMULVRSAVLINGPFTLNCD 547
 Db 402 FLGYSGARDIEGENELPRLVYVSREKPGYQOHKKAGAENALVRSVLTINAPFLNCD 461
 Qy 548 HYNNSKALRACMFLMDNLGKOVYVOPORDGIDKNDRAKRNNTVFEDVNMKLG 607
 Db 462 HYNNSKAVREAMFLMPQVGROCYVQFOPORDGIDORSDRYANRNTVFEDVNMKLG 521
 Qy 608 IOGRVYVGCVENRNTALGYEPPIKVKHKPKPSLISKLGCSRKRKSKAKESDKKSCR 667
 Db 522 IOGRVYVGCVENRNTALGYEPPIKVKHKPKPSLISKLGCSRKRKSKAKESDKKSCR 607
 Qy 668 HTSTVPUFNLDDIEEGVEGAGEDD-EKALLMSQMSLERKFGOSAVFVASTLMENG 726
 Db 580 REELDAAFNLREID-----NVEDYERSMISOTSFEKTFLGSSLVTESTLMENGVAE 633
 Db 727 SAPENILKEAHTHVISCGEDKSWMGMEIGWIKGSVTEIDLITGKMHARGWRSYCMPL 786
 Db 634 SAPSTPLIKEAHTHVISCGYECFKAWKGELKGWIGVGSVTEIDLITGKMHARGWRSYCMPL 693
 Qy 787 PAFKGSAPINSLSDLRNLQVRLWALGSVETLFSRHCPIWYGY-NGRKFLERFAYNTTI 845
 Db 694 PAFKGSAPINSLSDLRNLQVRLWALGSVETLFSRHCPIWYGY-NGRKFLERFAYNTTI 753
 Qy 846 ITSPLIMCTLLAVCLFTNOFTIPOISNIASWFLSFLSIFTANGLEMRSVCGIDW 905
 Db 754 FTSLPLIYCVBACILTGKFPIESENFASTWFLSIFTANGLEMRSVCGIDW 813
 Qy 906 WREOFWVIGGYSAHLRAFVFOGTLKVLAGIDTNTWKAASDEGDFDALLYLKWTLL 965
 Db 814 WREOFWVIGGYSAHLRAFVFOGTLKVLAGIDTNTWKAAD-DADFGLYVWTLL 872
 Qy 966 PPPTTLLIVLNWVGWAGSYAINGYQSMPLGKLFFAEWVWVLYPELKGLMQRNR 1025
 Db 873 PPPTTLLIVLNWVGWAGSYAINGYQSMPLGKLFFAEWVWVLYPELKGLMQRNR 932
 Qy 1026 TPTIVVVWSULLASFLSLWVRRIDFTSRVGPDTLE--CGINC 1065
 Db 933 TIVVLLWSVLSVASYLSLWVNPFTVSDTWSQSCISIDC 974

PRIOR FILING DATE: 1996-10-29
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 7
 LENGTH: 685
 TYPE: PRT
 ORGANISM: *cossipium hirsutum*
 US-08-960-048-7

Query Match 49.4%; Score 2803; DB 4; Length 685;
 Best Local Similarity 74.5%; Pred. No. 5.9e-276; Mismatches 76; Indels 22; Gaps 6;
 Matches 512; Conservative 77; MisMatches 76; Indels 22; Gaps 6;

Qy 396 ARKKWPFCKKVSIERAPAEWFAAKIDYLKDVKVQTSFKVKDRAMREYEETKIRINALVS 455
 ||:|||||:::|||||:||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 1 ARRWPFCKKHNRVERAPAEVFNEKIDLKDKVHFSFKERRAMREYEETKIRINALVA 60
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Qy 456 KALKCPEEGWMQDGTPWPGNNTDHPGMIQYLVGSAGALUDGKELPLRVYSRKRPG 515
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 61 KAQQKPEEGWMQDGTPWPGNNTDHPGMIQYLVGSAGALUDGKELPLRVYSRKRPG 120
 ||:|||||:|||||:|||||:|||||:|||||:
 Qy 516 FOHKKAGAMMALVRYSAVLUTNGPFLNLQDHYINNSKALREAMCFLMPNLGKOVCVY 575
 ||:|||||:|||||:|||||:|||||:|||||:
 Db 121 YOHKKAGAENALVRYSAVLUTNAPFLNLQDHYINNSKANREAMCFLMPNLGKOVCVY 180
 ||:|||||:|||||:|||||:
 Qy 576 QPPQRDGIDKNDRYANRNTVFFDINLRGLDQIOPGVVVGTVGCVFNRITALGYEPPIKV 635
 ||:|||||:|||||:|||||:
 Db 181 QFPQREDGIDPHDRYANRNVFFDINMLGLDGLQGVYVGTVGCVFNRQALGYDPVSEK 240
 ||:|||||:|||||:
 Qy 636 HKK-----PSILSKLGGSRKNSKAKE-----
 ||:|||||:|||||:
 Db 241 RPKWMTCDWPSCWCCCCGSGSRKSKKGEEKGGLLGGKKKKMKGKVVKGSSAPVFD 300
 ||:|||||:|||||:
 Qy 678 LDDIEGVGEGAFDD-EKALLMSQMSLEKRGQSAYFVASLMEENGPPSATPENLKE 736
 ||:|||||:|||||:
 Db 301 LEEIEEGLE--GYERLEKSTMSQNKFEKRGQSPVIASTLMEENGPPCTNSTLKE 358
 ||:|||||:|||||:
 Qy 737 AHNVSCGYEDKSDMWEIGMIGSYTEDILTGFKHARGRSITCCKMPKLAFKGSAPIN 796
 ||:|||||:|||||:
 Db 359 AIHVSCGYEKTEKNGKEIWIYGVSYTEDILTGFKHARGRSVCKPKRAFKGSAPIN 418
 ||:|||||:|||||:
 Qy 797 LSDRLHQVLRALGSAEVLFSRHCPIIWYGYNGRLKFLERFAYVNTIYPITSIPLMYCT 856
 ||:|||||:|||||:
 Db 419 LSDRLHQVLRALGSAEVLFSRHCPIIWYGYGGKWLDERLAYINTIVPFTSIPLAYCT 478
 ||:|||||:|||||:
 Qy 857 LLAVCLIFTQNTIPOISNLSIWIIFLSLSFATGLEMRSWGVGDEWMRNEQFWIGG 916
 ||:|||||:|||||:
 Db 479 IPAVCILTGKLIPTSLNTSWSWFLAFLLSTIATGTYLERLWGVSIODWRNEQFWIGG 538
 ||:|||||:|||||:
 Qy 917 VSAHLEAVFOGLVKVLAGIDNTFTSKASDEDGDFEAELYLFKWTTLLIPPTTLIUNLV 976
 ||:|||||:|||||:
 Db 539 VSAHLEAVFOGLVKVLAGIDNTFTSKASDEDGDFEAELYLFKWTTLLIPPTTLIUNLV 597
 ||:|||||:|||||:
 Qy 977 GVVAGSYAINSgyGSwgPLRGLKPAFWVIVHLVLPFLKGLGMGRONTPTIVVWSVLLA 1036
 ||:|||||:|||||:
 Db 598 GVVAGSYAINDGYGSwgPLRGLKPAFWVIVHLVLPFLKGLGMGRONTPTIVVWSVLLA 657
 ||:|||||:|||||:
 Qy 1037 SIFSLUWVRIDPFTSAYVGPDILEGI 1063
 ||:|||||:
 Db 658 SIFSLUWVRIDPFTSAYVGPDILEGI 684
 ||:|||||:
 RESULT 4

US-08-960-048-11
 Sequence 11, Application US/08960048C
 Patent No. 6271443
 GENERAL INFORMATION:
 APPLICANT: Staker, D. et al.
 TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
 TITLE OF INVENTION: Sequences
 FILE REFERENCE: 15621/01/US
 CURRENT APPLICATION NUMBER: US/08/960,048C
 CURRENT FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/029, 987
 PRIOR FILING DATE: 1996-10-29
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 11
 LENGTH: 693
 TYPE: PRT
 ORGANISM: *Escherichia coli*
 US-08-960-048-11

Query Match 4.9%; Score 278; DB 4; Length 693;
 Best Local Similarity 19.3%; Pred. No. 7.7e-9; Mismatches 191; Indels 362; Gaps 30;
 Matches 153; Conservative 86; MisMatches 191; Indels 362; Gaps 30;

Qy 260 MYIMMRUVICLICFLHYRIT-----APVNPALMIVSVCEIWALSWIDQPFKWPV 313
 ||:|||||:|||||:
 Db 21 MJIVLSSLJIVSCRYIWRYTSTLNWDDPVSLVCGILLFAITYAWTIVL-VLGYFQVWPL 78
 ||:|||||:
 Qy 314 NRETYLDRLALRDREGEPSQLAQANDVFDIYSTDPUKEPLPVANTVSLAVIDYDVKVS 373
 ||:|||||:
 Db 79 NRQP---VPJPKMSLWPS---VDIFVPTN---EDLNWVKTNTIASLGIDWKDKL 127
 ||:|||||:
 Qy 374 CVFDGGAAMSFSLEAETSEFARKWVPCFKYSIESRAPEWFAAKIDYLKDVKQTSFV 433
 ||:|||||:
 Db 128 IWILDDGG-----
 ||:|||||:
 Qy 434 KDRRAMREYEFKIRINALVSKALCPEEGWMQDGTPWPGNNTGDHPGMLQVFLGONG 493
 ||:|||||:
 Db 136 -----REFROF-----
 ||:|||||:
 Qy 494 GLDAEGNELPRLYVVSREKRFQFHKKAGAMNALVRSAVLTNGPFLNLCDHVINS 553
 ||:|||||:
 Db 147 GVK-----YIARUT---HEAKAGNTNNAKY---KGEEVSIIDCDHVPTRS 188
 ||:|||||:
 Qy 554 KALREAMCFLMDPNLGKQVYQFQPRE--DGIDKN---DRYANRNTVFRDINLRGLD 606
 ||:|||||:
 Db 189 FIQTMOWFELKE---KOLAMMQTPHFFSPDPFERNLGRFRKTPNGETLTQYGLVQDGD 244
 ||:|||||:
 Qy 607 GIQGPVYVGTGCVFNRITALGYEPPIKVHKPKPSLASKLCGGSRKNSKAKKESDRKKSG 666
 ||:|||||:
 Db 245 MWDTAFFCGSCAVIRR-----
 ||:|||||:
 Qy 667 RHTDSTVVFNFNLDIEGVVEGAGPDDBKALMSQMSLEKRGQSAVFASTLMEENGVPP 726
 ||:|||||:
 Db 263 -----LDEI-----
 ||:|||||:
 Qy 727 SATPENILKEATHVSCGYEDKSDMWEIGMIGSYTEDILTGFKHARGRSVTCCKMPK 786
 ||:|||||:
 Db 270 -----AVE-----
 ||:|||||:
 Qy 787 PAFKGSAPINSLSDRNLQVLRWALGSAEVLFSRHCPIIWYGYNGR-LKFLERFAYVNTIYP 845
 ||:|||||:
 Db 297 PQAAGLATESLSAHIGORIRWARGMWQI-FRDLNPL---TGKGLKAQORLQCYVNMHF 351
 ||:|||||:
 Qy 846 ITSPILLMYCT-----LLAVCLIFTQNPFIQI-----SNIASIWFLSFLS 886
 ||:|||||:
 Db 352 LSIGPRLFLTAPLAFLHAYIIYAPALMIALFVPHMITHASLTNSKIOCKYRHSFWE 411
 ||:|||||:
 Qy 887 IFAT-----GLEMRSWGVGDEWMRNEQF--WWIGGVSASHFAVE 925
 ||:|||||:
 Db 412 IYETLAWYIAPTLVALINPHKGKFNTANGGGLVE---EEYVWVVIS--RPYFLFLV 465
 ||:|||||:
 Qy 926 OGILKVLAGIDNTFTSKASDEDGDFEAELYLFKWTTLLIPPTTLIUNLV 976
 ||:|||||:
 Db 466 LNLVGVAVGI-----
 ||:|||||:
 Qy 977 GVVAGSYAINS 988
 ||:|||||:
 Db 502 VLGGAVAVSVS 513
 ||:|||||:

RESULT 5

US-08-960-048-12

```

Db 635 ARLESLSRPFVIFALLLVAMAFV 659
RESULT 6
US-08-960-048-10
; Sequence 10, Application US/08960048C
; Patent No. 6,071,443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: plant cellulose Synthase and Promoter
; TITLE OF INVENTION: sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960, 048C
; CURRENT FILING DATE: 1997-10-29
; PRIORITY APPLICATION NUMBER: 60/029, 987
; PRIORITY FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 756
; TYPE: PRT
; ORGANISM: acetobacter xylinum
; US-08-960-048-10

Query Match          4.2%; Score 236.5; DB 4; Length 756;
Best Local Similarity 19.3%; Pred. No. 1.5e-14;
Matches 166; Conservative 98; Mismatches 263; Indels 331; Gaps 34;
Oy
204 ERNTGPVSTQASERGGDIDASTD---ILADEALINDEAR----OPL 244
Db 11 ESRLGRISNKIILSRLGASYIVGALGLCALIAATTVTINNQOLIVAAVCVVIFFVWRGK 70
Oy
245 SRRKSITPSSRNPYRMVIMRLVLICLFLHYRITNPVNAFALWL-----VSVCBIEW 298
71 SRRTQI-----FLEVLSALVSLRYLTWRLTELD--FNTWIOGILGVILLAMELYA 119
Oy
299 ALSWILDQFPKWPVNRETYLDLRLALRYDREGESQLAVIDFVSVDPDLPKEPPVANT 358
Db 120 LYMLFLSYFOTIOPHLRAP---LPLPDNVDDWPT---VDIFIPTD--EOLSTVRLT 168
Oy
359 VLSILAVDYPVDKVSCVVFDDGAAMLFSFESLAETSEFARKWWVFPCKYSIEPRAPEWFA 418
Db 169 VLGALGIDWPDPDKVNVYLDDG-----VLSILAVDYPVDKVSCVVFDDGAAMLFSFESLAETSEFARKWWVFPCKYSIEPRAPEWFA 418
Oy
419 AKIDYLDKDVQJTSFKVKDRRAKMREYEFKIRINALSKALKCPEEGWVMDGCTPWGPNNT 478
Db 191 -----VRPEFEOF-----AKDC----- 202
Oy
479 GDHPGMIQVFLGQNGGLDAEGENELPRLVIVSREKRCFGHQHKAGAMNLRVSAVLTG 538
Db 203 -----GALYI-----GRVD-----SARAKAGNMHNALK----RTSG 231
Oy
539 PFTLNIDCDHYTNNSKALREANCFLMPNGLKQVWVQFOPRFDGIDKNDRYANRNFTVF 598
Db 232 DYILLDCDH IPTRAELOQIANGMWV--ADRKLALMQTPHFFYSDP----- 275
Oy
599 DINLRGLDGIGQGPVWVGTGCYFNRTALGYBPIKVKKHKKPSLLSKLCGGSRKKSKAKK 658
Db 276 -----FORNLAVGYRTP----- 287
Oy
659 ESDKKKSGRHDTSTVVFNLDDIEVGEGAGEDDEKALLQMOSLERKRGOSAVFVASTL 718
Db 288 -----PFGNL---FYFGVQDND-----FWATTEFGSCA 314
Oy
779 SIYCMPLKPAFKGSGAPINLSDRNQVRLWALGSVTEILFSRHCPIWVYNGR-LKFLERF 837
Oy
719 MENGGVPPSATPENLLEKAIHVISCGYEDKSDWMGEIGWIYGSVTDLTGFKMHARGWR 778
Db 349 TAYL--RIPVAGSLATERLTHIGORMWRARGMIOI-FRVNPNML---GRGKLGQRLC 401
Db 315 I-----LREATESIG-----GEAVETVDEAHTALMRQRGWG 348

```

Db 402 YLSAMTSFFFAIPRVLFLASPLAFLFAGONTIAAPLAVAYALPHMFHSTAT----- 454
 Qy 898 SGVGDIEWNNEQFW-VIGGVSAAHLAEVFOCILKLAGIDNTFTWSKA--SDEGDDFA 953
 Db 455 -AAVNKGWR-YSESVETMAFLVRVTIVTLEPSKGKNTYEGVLEEEFDLG 512
 Qy 954 ELY----LFWTTLIIPPTLLI-VNLVGVVAGVSYATNSGYOSWGLPFLGKLFCAFNV 1006
 Db 513 ATYPNLTIFATIMGGGLIGLFLELIVRNQLOVIARNAYLNCA-----WA 557
 Qy 1007 IVHLPFLKGL-MGRQR 1023
 Db 558 LISLILFAAIAVGRET 575

RESULT 7
 Sequence 3, Application US/09147236A
 ; Parent No. 6316251
 ; GENERAL INFORMATION:
 ; APPLICANT: TONOUCHI, Naoto
 ; APPLICANT: TSUCHIDA, Takayasu
 ; APPLICANT: YOSHINAGA, Fumihiro
 ; APPLICANT: TAHARA, Naoki
 ; APPLICANT: HAYASHI, Takahisa
 ; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
 ; FILING REFERENCE: 6537-011-OPTC
 ; CURRENT APPLICATION NUMBER: US/09/147,236A
 ; CURRENT FILING DATE: 1999-04-08
 ; EARLIER FILING NUMBER: PCT/JP97/03653
 ; EARLIER FILING DATE: 1997-10-09
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 756
 ; TYPE: PRT
 ; ORGANISM: Acetobacter xylinum
 ; FEATURE:
 ; OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, t, or
 ; OTHER INFORMATION: c
 ; US-09-147-236-3

Query Match 4 2%; Score 236.5; DB 4; Length 756;
 Local Similarity 19.3%; Pred. No. 1.5e-14;
 Matches 166; Conservative 98; Mismatches 263; Indels 331; Gaps 34;
 Qy 204 EKNTGPVSTQASERGGVIDDASTD---ILADEALLNDEAR-----OPL 244
 Db 11 ESRLRGRISKNLKSLRGASYIVGALGICLALIAATTWLNNEBOLTVAAVCVWIFFVNGRK 70
 Qy 245 SRKVSPSPSRINPYVMVIRLWVICLEHYRITNPVPAFLWL----VSVICEIW 298
 Db 71 SRRTOQI-----FLEVSLVALSLRYLWRTELD-FNTWLOGIQLVILMELVA 119
 Qy 299 ALSWILDQFPKWFPPVNRTEYLDRALARVREGEPSOLAVIDFVSTDPLKEPPVLTANT 358
 Db 120 LYMLFLSYFGTIOPLHRAP---LPDPDNYDWP---VDFIPIYD---EQLSIVRLT 168
 Qy 359 VLSLAVDYPDVKSYCYYFDGGAAMLSEFESLAETSEFARKWNPCKKYSIEPRAEWYFA 418
 Db 169 VLGALGIDWPPDKVNVYLDDG----- 190
 Qy 419 AKIDYIKDKVQTSFKDRRAMREYEEFKIRINALVSKALKCPEEGWWMDGTPWPGNNT 478
 Db 191 -----VREPEREOF-----AKDC----- 202
 Qy 479 GDHPGMIQVFLGQNGGLDAECDNELLPRLVYVSREKRGFQHKKAGAMNLVRSVAVLNG 538
 Db 203 ---GALYI----GRVVD-----SAHAKAGNLNAIK---RTSG 231
 Qy 539 PFILNLDCKDHYINNSKALRACMCFIMDPNGKQCVQFCVQRFDGTDKNDRYANRNTVFF 598

RESULT 8
 Sequence 2, Application US/07689008
 ; Parent No. 5268274
 ; GENERAL INFORMATION:
 ; APPLICANT: Ben-Bassat, Arie
 ; APPLICANT: Calhoon, Roger D
 ; APPLICANT: Fear, Anna L
 ; APPLICANT: Gefland, David H
 ; APPLICANT: Meade, James H
 ; APPLICANT: Tal, Ronny
 ; APPLICANT: Wong, Hing
 ; APPLICANT: Ben-Ziman, Moshe
 ; TITLE OF INVENTION: METHODS AND NUCLEIC ACID SEQUENCES FOR THE
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: McCutchen, Doyle, Brown & Enersen
 ; STREET: Three Embarcadero Center
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MC-DOS
 ; SOFTWARE: Patentin Release #1.24
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/689, 008
 ; FILING DATE: 19910422
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 337, 194
 ; FILING DATE: 12-APR-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 496, 236
 ; FILING DATE: 23-MAR-1990

Db 232 DVILILDCDH-IPTRAFLOIAGMWY--ADRKIALMOTPHFYSPDP----- 275
 Qy 599 DINLRGLDGIGPVPVYGTGCYFNRTALYGEPPPIKVHKKSSLSSLCGGSRKNSKAKK 658
 Db 276 -----FORNLAVGVRP----- 287
 Qy 659 ESDKKSGRHTDSTPVFNLDIEGVEGAGFDDEKALLMSQMSLKEKGOSAVFVASTL 718
 Db 288 -----FBGNL---FQVYIQCND-----FMDATPCGS 314
 Qy 719 MENGGVPPSATPENLIREAIIHVISCGYEDKSNDWGMBIGWIVGSVTDILNGFKMWHGWR 778
 Db 315 I-----LRKATESIG-----GFAVETVTAHDALTALMRQRGWS 348
 Qy 779 SIYCMPLKPAFKGAPSAPINSURNQVLRWALGSVETLFSRCPWVYNGR-LKLERFA 837
 Db 349 TAYL--RIPVASGLATERLTHIGORMWRARGMIQT-FRVNPNML---GRGLKLQRLC 401
 Qy 838 YVNTTIPITSIPLMCTIILAVCLFTQNFIPOQSNIASTWFLSPLSISATGILEMRW 897
 Db 402 YLSAMTSFFAIPRVLFLASPLAFLFAGONTIAAPLAVAYALPHMFHSTAT----- 454
 Qy 898 SGVGDIEWNNEQFW-VIGGVSAAHLAEVFOCILKLAGIDNTFTWSKA--SDEGDDFA 953
 Db 455 -AAVNKGWR-YSESVETMAFLVRVTIVTLEPSKGKNTYEGVLEEEFDLG 512
 Qy 954 ELY----LFWTTLIIPPTLLI-VNLVGVVAGVSYATNSGYOSWGLPFLGKLFCAFNV 1006
 Db 513 ATYPNLTIFATIMGGGLIGLFLELIVRNQLOVIARNAYLNCA-----WA 557
 Qy 1007 IVHLPFLKGL-MGRQR 1023
 Db 558 LISLILFAAIAVGRET 575

ATTORNEY/AGENT INFORMATION:
 NAME: Murphy, Lisabeth Felix
 REGISTRATION NUMBER: 31547
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 393-2286
 TELEFAX: (415) 393-2200
 TELERX: 340817 MACPAG SFO
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3031 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-689-008-2

Query Match 4.0%; Score 228.5; DB 1; Length 3031;
 Best Local Similarity 19.3%; Pred. No. 1.2e-12; Matches 154; Conservative 84; Mismatches 229; Indels 329; Gaps 30; QY 263 MRLWVILCFLHYRITNPVNAFLWL----VSVICWFLSWLDOOPPKWEPVNRE 316
 QY 80 VLSALWSLRVLTYLWRIETLD-FDTWIQGGLVILLMAELYALMFLSYQTQPLHRA 137

317 TYLRLALRYDREGEPSQLAAVDIFVSTVPLKEPLPVANTVLSILA 138 P---LPLPDNVDDWPT---VDIFIPTYD--BOLISIVLTVLGALGIDWPPDKRVNVI 186
 377 FDDGAMLSFESLAETSEFARKWVPFCKKYSIEPRAPENWFAAKIDYLKDVKQTSFVKDR 436
 187 LDDG----- 190

437 RAMKREYEEFKIRINALVSKALKCPEEGWVMQDGTPWPWGNNTGDHPGMICVFLGGGLD 496
 191 -VPEFEQPAKDGCGA----- 205

497 AEGNELLPRLVVYSBKKRGFQHHKAGAMNALVRVSAVLTINGPFLNLDCDHYINNSKAL 556
 QY 206 -----YIGRVD---SSHAKAGNLNHAIA---RTSGDYIILPDCDH-IPTRAFI 246
 557 REAMCFLMDPNLKGQCVYQFOPRFIDKNDRYANRNTVFFDINLRGLDGIQGPVVYGT 616
 247 QIANGWMV--ADRKIALMOTPHFVSPP----- 273

617 GCVNENTRALYGYEPPKVKHKKPSSLKLCGGSRRKNSKAKKESDKKKSGRHPTSTVPVF 676
 274 --FORNLAVGYRTP----- 288

677 NLDDIECGVEGAGDDDEKALLMSQMSLEKRGFGOSAVFVASTLMEENGGVPSATPENLKE 736
 QY 289 NL---FYGVTDGND-----FWDATFCCSCAI-----LRRE 317

737 AIHVISCGYEDKSDWGMGEIWIYGSVTEIDLTKRMHARWRSTIYCMPKLPFAFKGSSAPIN 796
 318 AIESIG-----GFAVEITVDETAHTLMQRGWSSTAYL-RIPVASGLATER 362

797 LSDRNUQVIRWALSVEILFSRHCPIWGYGRKFLLERAYVNTTIPITSIPILMYCT 856
 QY 363 LITHIGORMRWARGMII-FRVDNFM--LGGGLKQRLCYLSAMTSFFAIRVIFLA 418

857 LLAVCLFTNOFLIQOISNTIASIWLSLFLSIFATGLEMRS--GVGIDEMWRNEQFW- 912
 QY 419 SPLAFLFFGONTIASP-----LAVLAYAIPHMFHSIATAKYKNGWR-YSFNS 466
 Db 970 LLIVLNVLGVVAGVSYAINGQSYQSPFLG-KLFAFWVIVHLYPFLKGLMGGRNRPPTI 1027
 913 VIGGSVAHIFAVFOGILKVIAGIDNTNTFTSKA--SDEGDFABLYLFKWTTLIIPPT 969
 467 EVVTTMALELVRVTITLMPSKGKFNWPKGGVLEEEFDGATY-----PNI 516

QY 1028 VVWS-----VLLASI 1038

RESULT 9
 US-08-960-048-9
 Sequence 9, Application US/08960048C
 Patent No. 6271443

CURRENT FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/029, 987
 PRIORITY FILING DATE: 1996-10-29
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 9
 LENGTH: 723
 TYPE: PRT
 ORGANISM: Agrobacter xylinum
 US-08-960-048-9

Query Match 3.5%; Score 196; DB 4; Length 723;
 Best Local Similarity 20.1%; Pred. No. 1.9e-10; Matches 161; Conservative 93; Mismatches 241; Indels 306; Gaps 36; QY 251 PSRINPRYMMVILRLVILCFLHYRITNPVNAFLWL----VSVICWFLSWL 304
 QY 70 PRRSQIF-LEVLSQLVSL-RVLTWRITEL--SFDTWLOGLGMVLAELYALMFL 125
 Db 305 DQFPKNEPVNRETYLRLALRYDREGEPSQLAAVDIFVSTVPLKEPLPVANTVLSILA 364
 Db 126 SYFOTIAPLHRA-----LPL----PPNPDEWTFDVIFVPTN--EELSTVRLTVLGSLG 174
 Db 365 VDYPVDRKVCVVFDDGAAMLSEFESLAETSEFARKWVPFCKKYSIEPRAPENWFAAKIDYL 424
 Db 175 IDWPPEKVVRHILDDG-----RRPE-FAA----- 197

425 KDKVOTSFVKDRRAMKREYEEFKIRINALVSKALKCPEEGWVMQDGTPWPWGNNTGDHPGM 484
 Db 198 -----FAEC----- 202

485 IOVFLGONGGLDAEGNELPRLVVYSBKKRGFQHHKAGAMNALVRVSAVLTINGPFLNL 544
 Db 203 -----GAN-----YIARP-----NEHAKAGNLYAIG---RTGDDYLIF 235

545 DCDDHYINNSKAKLREAMCFLM-DPNLKGQCVVQFOPRFIDGIDKNDRYANRNTVFFDINLR 603
 Db 236 DCDDH-VPTRATIQLTQGMWEDP---KIALMQTPRHFVSPP----- 273

604 GLDGIQGPVVYVGTCVFNRTALYGYEPPKVKHKKPSSLKLCGGSRRKNSKAKKESDK 663
 Db 274 -----FORNLASGYRTP----- 285

664 KSGRHTDSTPVWNLDIECGVAGCDFDEKALLMSQMSLEKRGFGOSAVFVASTLMEENG 723
 Db 286 -----REGNL---FYGWQDGND-----FMDATEFCGSCAI---- 313

724 VPPSATPENLKEIAHVISCGYEDKSDWGMGEIWIYGSVTEIDLTKRMHARWRSTIYC 783
 QY 314 -----IRRATIAEQIG-----GFATQTVFEDAKTALKMQRGWSSTAYL- 350

784 PKLPAFKGSAMTINLSPRNQVRLWALGSVEILFSRHCPIWGYGR-LKTFLERAYVNTT 842
 Db 351 -RPLLAGLATERLTHIGQVRWARGMLQ-FRIDNPLF---GGLSLMQRCLCYLSM 404

843 IIPITSIPLMLCTLBACVLETOFIQOISNIAIWLSLFLSIFATGLEMRLMSGVG- 901
 Db 405 TSFLFAVPRVFLSLBLAFJFGQNTIAAS-----LALLAYAIPH-FHAUT 452

QY 902.----IDWWMRNEQFW-VIGGSVAHIFAVFOGILKVIAGIDNTNTFT\$KASDEGDDAELYL 957

RESULT 10
 PCT-US91-01726-4
 ; Sequence 4, Application PC/US9101726
 ; GENERAL INFORMATION:
 ; APPLICANT: Brown, Malcolm R
 ; APPLICANT: Saxena, Indra M
 ; TITLE OF INVENTION: Recombinant Cellulose Synthase
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David L. Parker
 ; STREET: 600 Congress Avenue Suite 2300
 ; CITY: Austin
 ; STATE: Texas
 ; ZIP: 78701
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PC/US91/01726
 ; FILING DATE: 19910314
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 494093
 ; FILING DATE: 15-MAR-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parker Mr, David L
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 512-320-7200
 ; TELEFAX: 512-474-7577
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 723 amino acids
 ; TYPE: AMINO ACID
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; PCT-US91-01726-4

Query Match 3.5%; Score 196; DB 5; Length 723;
 Best Local Similarity 20.1%; Pred. No. 1.9e-0; Matches 161; Conservative 93; Mismatches 241; Indels 306; Gaps 36;

Qy 251 PSSRNPYRMVIMRLUVICLFLHYRINTPVPNAFLWL-----VSVICEWFALSWIL 304
 Db 70 PSRRSQIF-LETVLSGLNLS-RYLTWRIETL-SEDIWNLQHGTMLVIAEYLALMFL 125

Qy 305 DOPPKWTFPVNRETYLDRLALRYDREGESQLAANDIVFSTVPLKEPPLVANTVSLA 364
 Db 126 SYFQTAPLHRAP----LPL---PPRDEPTVDFVPTY--ERLSIVLTVLGSIG 174

Qy 365 VDPVPDKSCVYFDGGAMLSFSELASTMSEFARKWYFCKKYSIEPRAPENWFAAKIDYL 424
 Db 175 IDWPPEKVRVHLDDG-----RRPE--FAA---- 197

Qy 425 KOKWQSFVKDRAMKEYEEFKIRINAVALSKALKCPEGWMQDGTPWPGNNTGHDHGM 484
 Db 198 -----FAAC----- 202

RESULT 11
 US-08-119-773-4
 ; Sequence 4, Application US/08119773
 ; Patent No. 5460942
 ; GENERAL INFORMATION:
 ; APPLICANT: Chou, Janice Y.
 ; APPLICANT: Lei, Ke-Jian
 ; APPLICANT: Shelly, Leslie L.
 ; TITLE OF INVENTION: GLUCOSE-6 PHOSPHATASE: THE GENE AND PROTEIN AND RELATED MUTATIONS
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Khourie and Crew
 ; STREET: Steuart Street Tower, One Market Plaza
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 94105-1493
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/119, 773
 ; FILING DATE: 10-SEP-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weier, Kenneth A.
 ; REGISTRATION NUMBER: 31,677
 ; REFERENCE/DOCKET NUMBER: 15280-175
 ; TELECOMMUNICATION INFORMATION:

TITLE OF INVENTION: GLUCOSE-6-PHOSPHATASE: THE GENE
 TITLE OF INVENTION: PROTEIN AND RELATED MUTATIONS
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Khourie and Crew
 STREET: Stewart Street Tower, One Market Plaza
 CITY: San Francisco
 STATE: California
 STATE: California
 COUNTRY: US
 ZIP: 94105-1493
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/1119,773
 APPLICATION NUMBER: US/08/1119,773
 FILING DATE: 10-SEP-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Weber, Kenneth A.
 REGISTRATION NUMBER: 31,677
 REFERENCE/DOCKET NUMBER: 15280-175
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 543-9600
 TELEFAX: (415) 543-5043
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 346 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 1..346
 OTHER INFORMATION: /label= Truncated prot
 OTHER INFORMATION: /note= "G-6-Pase truncated due to mutation at nucleotide 1118 codon 347"
 OTHER INFORMATION: to a stop codon at 347
 OTHER INFORMATION: 08-119-773-5
 Very Match 2.0% Score 115; DB 1; Length 19.5%; Pred. No. 0.0087; Mismatches 123; In
 matches 78; Conservative 44; Mismatches 123; In
 736 EAHVHSQGYEDKSDWGMGMEIGWIGVGSVTEDLTGFKMHARGWR
 :: | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 14 OSTHYLQVNQYQDSQDW---FILVSVIADLRNAYFVLFPIW-
 796 NLSDRMLQVWLWALGSGVEELFSRHCPIWG-----
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 66 VIGDWNLVLFKW-----ILFGQR-PYWWLDTDXSNTSVPLL
 826 -----YNGRLKFLEPRRAYVNNTIYIPITSI-----
 119 HAMGTAGVYVYVMTSLSIFQGKIKPTYRERCLNVILW-----
 863 -----FTMNOFTIPOTSNI-----STMFLSL-----F
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 173 LAAHFPHOVWAGVLISGLGIAVETFSIHSHINSASLKKYFLITFF
 902 IDEWWRE--QFWIGGYSAAHLFAVEQGTLKVLAGIDTINFVIT
 232 VDLWLTWLEKAQRMCEOPEWVH-----IDTT-----
 960 WTTLLIPTTLIVLNVLGVGVAGVSYAINSG-----YOSWW
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 257 -----PEASLKLNL-GTFGLGLALNNSMYRESCKGKLSKWW

RESULT 14
US-08-119-773-2
 sequence 2, Application US/08119773
 patent No. 5460942
GENERAL INFORMATION:
 APPLICANT: Chou, Janice Y.
 APPLICANT: Lei, Ke-Jian
 APPLICANT: Shelly, Leslie L.
TITLE OF INVENTION: GLUCOSE-6-PHOSPHATASE: THE GENE AND PROTEIN AND RELATED MUTATIONS
NUMBER OF INVENTION: 36
CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Khourie and Crew
 STREET: Stewart Street Tower, One Market Plaza
 CITY: San Francisco
 STATE: California
 COUNTRY: US
 ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/119,773
 FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
 NAME: Weber, Kenneth A.
 REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-175
TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 543-9600
 TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 357 amino acids
 TYPE: amino acid
 TOPOLOGY: Linear
 MOLCULE TYPE: protein
US-08-119-773-2

Query Match 2 0% Score 115; DB 1; Length 195; Pred. No. 0.0092; 123; Index
Best Local Similarity 19.5%; Matches 78; Conservative: 44; Mismatches 123; Indels 0
QY 736 EAIHWISCGYEDKSQMGMEIGWIGSVTEDILTGKMHARGWRSI
Db :: | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 14 QSTHQLQVNQDSDQW----FLMSVIAIDRNARYVLPFW----
Db : | || | : | : | : | : | : | : | : | : | : | : | : | : |
QY 796 NLSDRNUQNLWRALGSVTEILFSRSHRCPIWYG----
Db 66 VIGDNMLNLVFKW-----ILFGQR-PYWWVLDTDYNSNTSYPLIM
QY 826 -----YNGRLKFLERAYVNTIYIPITSIP-----
Db : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 119 HAMGTAGVYYVMVMTLSIFQGKPTYRFCLNLYLW-----
QY 863 -----FTNQFIIQPIQNSIA-----SFLWFLSL-----FLS-----
Db : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 173 LAAHPPHQVAVGVLSGIAYVETFSIHSIYNAISLKKYFLITFLE
Db 902 IDEWHRNE-OFWVIGGSVAHLFAVFOGILKVLACIDTNTFTVTSK
QY 232 VDLMLTLEKAQRCBQPEWVH-----IDTT-----
QY 960 WTTLLIPPTTLIVLNUVGVVAGVSAINSG-----YQSMGK

1 357;
2 als 154; Gaps 16;
3 YICMPKIPAFKGSAPI 795
4 : : |
5 FHLQEAVGKIKLWVA 65
6 ----- 825
7 QQPFPVTCETGPGSPSG 118
8 LMYCTTLLAVCL----- 862
9 : | | | |
10 GFYWAVQVNVLCSRIY 172
11 IFATGIGLEMMRWSGVG 901
12 : | | ; |:
13 SFAIGGY-LLIKGLG 231
14 ASDEBGGDFALYLFLK 959
15 LFGKJEFAPWVIVHL 1010
16 ----- 256

RESULT 15
US-09-000-094-46
Sequence 46, Application US/09000094
Patent No. 6365160
GENERAL INFORMATION:
APPLICANT: WERB, Elizabeth Ann
MARGETTS, Mary Brigid
COX, John Cooper
FRAZER, Ian
MCWILLIAN, Nigel Alan John
WILLIAMS, Mark Phillip
MOLONEY, Margaret Bridget
Holland
EDWARDS, Stirling John
SIREET, 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/000,094
FILING DATE: 21-Apr-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU96/00473
FILING DATE: 27-JUL-1996
APPLICATION NUMBER: AU PN 4439/95
FILING DATE: 27-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENT, St phen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 017227/0137
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1587 amino acids
TYPE: amino acid
TOPLOGI: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-000-094-46
Query Match 1.9%; Score 110; DB 4; Length 1587;
Best Local Similarity 19.3%; Pred. No. 0.44; Mismatches 114; Conservative 77; MisMatches 180; Gaps 30
Matches 2 ESEGETACKPMK--NTVPOTCOTCSDNVKGTVODDRFVACDICSPVCRCPYERKDG 58
Db 877 EVDGQDS-QPLKOHFOITCCCG-CDSNRVLVQ-----CTEDIREV 917
Qy 59 NO-----SCPOCKTRYKRLKGSPAIPGDKDDEGLADBGTVENPQKEKISRMGLW 110
Db 918 QOQLLGTLNIVCPIC-----APKTSMADDSTENEGSGCTGWMVEAIVQHPTGT 968
Qy 111 HLRKGKGEEMGERQY-----DKEVSHHL--PRILTSRDTSGFSASAERLSVSSTI 161

Db 969 QISDDDEBEVEDSGYDWFIDDSNITHNSLERAQALNRQEADTHY :|:
Qy 162 AGGKR---LPYSSDWNO-----SPNRITDPPGLGNWAK-----ERVD-G 198
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1018 ODLKRYKYGPSVSPINTIAEVESEISPR--LDAIKLTROPKVKRRLFOTRELDSG 1074
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Oy 199 WAKMOK-BKNTG-PVSTQASERG- -GUDIADSTIDLADEALLNDERAPOLSRKVSPVSSR 254
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1075 YGYSEVAGTGPQVEKEKGVPENGDDQEKDGDRDINGEE-----HTEAEAPTNS 1123
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy 255 .INPYRMVIMLRUICFLHYRTNPVNAFALWLSVI-----CEIWFAWSW-- 302
; ; :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1124 VREHAGSAGILELLIKEKDRLRALLGKRECEFCISFDLIRPKSDTICLDMWWVAGFGIH 1183
303 -ILDOFPKWF-PVNIRETYLDRLA-----LRYDREGEPSQA---AVDIFSTVD 346
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1184 HSISEAFOKLIEPLSLYAHIONLWTNAWMVLVLLRKFKVNRSTARTLATLNNPENQ 1243
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy 347 PIKEPFY-----TANVILSLAVDOPVDKSCIVDDGAMLSFF--- 387
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1244 MLIPEPKHOSGAVALWFRTGISNASTVIG---EARPEWITROTVEHGLADSOFLITEM 1299
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy 388 -----SLAETSEFARKWIFCKKSIEPRAPEWVFAAKIDYLKOKVQTSFVKDRRAMK 440
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1300 VOWAYDNDICESELA---FEVAQRFQEDFSNARA-----FLANSNOQARYVKOCATMC 1348
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Qy 441 REYEEFKIRINALVSKALKCPEEGWQNODGTPWPGNTGDPGMIOFELGN 492
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 1349 RHYKAEMR-----KMSIKQWIKHRGSKIEG--TGNWKPIVOFLRHON 1389
| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Search completed: June 16, 2003, 11:08:57
Job time : 25 secs

Page 1

SUMMARIES

SUMMARIES							RESULT 1
Result No.	Score	Query Match Length	DB ID	Description	PR	PD	XX
1	5655	99 6	1065 19	AAW3319	XX	08-JAN-1998.	XX
2	5654	99 6	1065 21	AAAG48133	PA	04-JUN-1997;	97WO-AU00402.
3	5654	99 6	1065 23	ABB93356	PA	24-JUN-1997;	97AU-0000699.
4	5597	98 6	1054 21	AAAG4734	PA	27-JUN-1996;	96AU-0000699.
5	5066	89 2	959 21	AAAG48135	(CSIR) COMMONWEALTH SCI & IND RES ORG.		
6	4572	80 5	1077 21	AAY84108	(AUSU) UNIV AUSTRALIAN NAT.		
7	4572	80 5	1077 21	AAY84114	PI		
8	4564	80 4	1079 23	AAU1096	Arioli A, Betzner AS, Peng L, Williamson DR		
9	4529	79 8	1076 21	AAY84121	DR		
10	4418	77 8	1043 21	AAY84119	N-PSDB; AAV06567.		

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

ALIGNMENT:

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:	June 16, 2003, 11:01:28 ;	Search time 45 Seconds
		(without alignments)
		3153.598 Million cell updates/sec
Title:	AAC93936	
Perfect score:	5677	
Sequence:	1 MSEEGETAGKPMKNIVPQTC.IDPFITSRVTGPDILECGINC	1065
Scoring table:	BLOSUM62	
	Gapp 10.0 , Gapext 0.5	
Searched:	908470 seqs, 1332505620 residues	
Total number of hits satisfying chosen parameters:	908470	
Minimum DB seq length:	0	
Maximum DB seq length:	200000000	
Post-processing:	Minimum Match 0%	
	Maximum Match 100%	
	Listing first 45 summaries	
Database :	A_Geneseq_101002:*	
	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1980.DAT:*	
1:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1981.DAT:*	
2:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1982.DAT:*	
3:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1983.DAT:*	
4:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1984.DAT:*	
5:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1985.DAT:*	
6:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1986.DAT:*	
7:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1987.DAT:*	
8:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1988.DAT:*	
9:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1989.DAT:*	
10:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1990.DAT:*	
11:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1991.DAT:*	
12:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1992.DAT:*	
13:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1993.DAT:*	
14:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1994.DAT:*	
15:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1995.DAT:*	
16:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1996.DAT:*	
17:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1997.DAT:*	
18:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1998.DAT:*	
19:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL1999.DAT:*	
20:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2000.DAT:*	
21:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2001.DAT:*	
22:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2002.DAT:*	
23:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2003.DAT:*	
24:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2004.DAT:*	
25:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2005.DAT:*	
26:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2006.DAT:*	
27:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2007.DAT:*	
28:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2008.DAT:*	
29:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2009.DAT:*	
30:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2010.DAT:*	
31:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2011.DAT:*	
32:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2012.DAT:*	
33:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2013.DAT:*	
34:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2014.DAT:*	
35:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2015.DAT:*	
36:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2016.DAT:*	
37:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2017.DAT:*	
38:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2018.DAT:*	
39:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2019.DAT:*	
40:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2020.DAT:*	
41:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2021.DAT:*	
42:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2022.DAT:*	
43:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2023.DAT:*	
44:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2024.DAT:*	
45:	/SIDS2/gcdata/geneseq/geneseqp-emb1/AAL2025.DAT:*	

xx
PS Claim 29; Page 144-149; 207pp; English.
xx
CC this polypeptide is encoded by cDNA clone Ath-B (see AAV06567)
CC that was isolated from an Arabidopsis thaliana cDNA library using
PCR primers (see NAT98632-34) based on cellulose synthase RSW1
CC genomic clone 23H12 (see AAV06563) and EST clone AAT20782 (see
AAV06562). It is closely related to Arabidopsis cellulose synthase RSW1
protein (see AAW3816-17). Claimed nucleic acid molecules (see
AAV06562-69) coding for claimed polypeptides (see AAW3816-20 and
AAW46202) involved in cellulose biosynthesis can be used to
manipulate the cellulose and/or beta-glucan content of transgenic
plants. Expression of nucleic acids in the sense orientation
increases the level of cellulose and reduces the level of
non-crystalline beta-1,4-glucan and starch, providing plants with
modified strength and/or shape and/or fibre properties, or having
increased resistance to stresses or pests. Antisense, ribozyme or
co-suppression molecules can be used to reduce the cellulose
content of a transgenic plant, e.g. to improve digestibility or to
alter carbon partitioning such that increased carbon is available
for growth, rather than deposited as cellulose.

SQ Sequence 1065 AA;

Query Match Local Similarity 99.6%; Pred. No: 0; Mismatches 3; Indels 0; Gaps 0; Matches 1062; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MESEGETAGKPMNIVPOTCQICSDNYGKTVDDREVAACIDCSFPVCPCYEVERKGQ 60

Db 1 MESGEETAGKPMNIVPOTCQICSDNYGKTVDDREVAACIDCSFPVCPCYEVERKGQ 60

QY 61 SCPOCKTRYKRKGSPATPGDKDKEGLADEGTEFENYFOKEKTSERMGWHLTRKGEM 120

Db 61 SCPOCKTRYKRKGSPATPGDKDKEGLADEGTEFENYFOKEKTSERMGWHLTRKGEM 120

QY 121 GEQYDKWSHNHLPLRSTQDTSGERASPERLSVSSPLAAGRLLPSSDVNQSPQR 180

Db 121 GEQYDKWSHNHLPLRSTQDTSGERASPERLSVSSPLAAGRLLPSSDVNQSPQR 180

QY 181 IVPVGVLGNWAWERVKDWMKQEKNGTGPVSTOASERGVGDIDASTDLABALLDEA 240

Db 181 IVPVGVLGNWAWERVKDWMKQEKNGTGPVSTOASERGVGDIDASTDLABALLDEA 240

QY 241 RQLRSRKVSPSSRINPYRVMILRVLICLHYRITNPVNAFLMLVSVICIEWFAL 300

Db 241 RQLRSRKVSPSSRINPYRVMILRVLICLHYRITNPVNAFLMLVSVICIEWFAL 300

QY 301 SWILDQFPKWPNNRETIIDLRLALRYDREGEPQLAANDIFVSTDPIKEPPVANTNL 360

Db 301 SWILDQFPKWPNNRETIIDLRLALRYDREGEPQLAANDIFVSTDPIKEPPVANTNL 360

QY 361 SILAVDYPDKVSCYVFDDGAMLFSESLAETSEFARKWPKRCKSYSTEPRAPEWYRAK 420

Db 361 SILAVDYPDKVSCYVFDDGAMLFSESLAETSEFARKWPKRCKSYSTEPRAPEWYRAK 420

QY 421 IDYLKDQVTSFYKDRRAMKREVEEKFIRKRNALVSKALKCPEEGWVMDGTWPGNNGD 480

Db 421 IDYLKDQVTSFYKDRRAMKREVEEKFIRKRNALVSKALKCPEEGWVMDGTWPGNNGD 480

QY 481 HPMQIQLFGQGLDAEGNELEPRLVVISERKPGFHKKAGMNLVRVSAVLNGP 540

Db 481 HPMQIQLFGQGLDAEGNELEPRLVVISERKPGFHKKAGMNLVRVSAVLNGP 540

QY 541 ILMEDCDHYINNSKALRAMECFMLDPNIGKQCYVQFRPFDICDKNDRYANRNTVFDI 600

Db 541 ILMEDCDHYINNSKALRAMECFMLDPNIGKQCYVQFRPFDICDKNDRYANRNTVFDI 600

QY 541 ILMEDCDHYINNSKALRAMECFMLDPNIGKQCYVQFRPFDICDKNDRYANRNTVFDI 600

Db 541 ILMEDCDHYINNSKALRAMECFMLDPNIGKQCYVQFRPFDICDKNDRYANRNTVFDI 600

QY 601 NURGLDGQGPVYVGTCYFNRLALYGYEPPTKVHKKPSLUSKLCGSRKNSKAKKES 660

Db 601 NURGLDGQGPVYVGTCYFNRLALYGYEPPTKVHKKPSLUSKLCGSRKNSKAKKES 660

QY 661 DKKSGRHTDSTVVFNLDDIEQVEAGGFDEKALIMQSLEKRGQSAVFASTIME 720

Db 661 DKKSGRHTDSTVVFNLDDIEQVEAGGFDEKALIMQSLEKRGQSAVFASTIME 720

QY 721 NGGTPSAPENLKEATHVISCGEDKSDWGMIGWLYGSWEIDLTGKFMARGWASI 780

Db 721 NGGTPSAPENLKEATHVISCGEDKSDWGMIGWLYGSWEIDLTGKFMARGWASI 780

QY 781 YCMRKKPAFKGSAPINLSDLNQVLWALGSVETLFSRCPWYGRKLEFERYAVN 840

Db 781 YCMRKKPAFKGSAPINLSDLNQVLWALGSVETLFSRCPWYGRKLEFERYAVN 840

QY 841 TTIVPITSIPLLMYCTLAUCLETNQFTIPOISNIASTWFLSFLSIFATGILEMMRNSGV 900

Db 841 TTIVPITSIPLLMYCTLAUCLETNQFTIPOISNIASTWFLSFLSIFATGILEMMRNSGV 900

QY 901 GIDMKWNRNQFWYVGVSAHLAEVFGTILKVLAGIDTNTVTKASASDGDRVELYLKWK 960

Db 901 GIDMKWNRNQFWYVGVSAHLAEVFGTILKVLAGIDTNTVTKASASDGDRVELYLKWK 960

QY 961 TLLIPPTILLVNWGVAGVSAYAINGQYSQNSQPLGKLFAAFWVIVHLPKLGLMR 1020

Db 961 TLLIPPTILLVNWGVAGVSAYAINGQYSQNSQPLGKLFAAFWVIVHLPKLGLMR 1020

QY 1021 QNRRTIVVWSVILLASIFSLWNRIDPFTRVYGPDILEGINC 1065

Db 1021 QNRRTIVVWSVILLASIFSLWNRIDPFTRVYGPDILEGINC 1065

RESULT 2

AAG48733

ID AAG48733 standard; Protein: 1065 AA.

XX

AAG48733;

XX

18-OCT-2000 (first entry)

DT

XX

DE

Arabidopsis thaliana protein fragment SEQ ID NO: 61573.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX

OS

Arabidopsis thaliana.

XX

EP10334-05-A2.

XX

PD 06-SEP-2000.

XX

PR

25-FEB-2000; 2000BP-0301439.

XX

PR 25-FEB-1999; 990US-0121825.

PR 03-MAR-1999; 990US-0123180.

PR 03-MAR-1999; 990US-0123548.

PR 23-MAR-1999; 990US-0125788.

PR 23-MAR-1999; 990US-0126264.

PR 23-MAR-1999; 990US-0126785.

PR 01-APR-1999; 990US-012462.

PR 08-APR-1999; 990US-0128234.

PR 08-APR-1999; 990US-0128714.

PR 16-APR-1999; 990US-0129845.

PR 19-APR-1999; 990US-0130077.

PR 21-APR-1999; 990US-0130449.

PR 23-APR-1999; 990US-0130510.

PR 23-APR-1999; 990US-0130891.

PR 28-APR-1999; 990US-0131449.

PR 30-APR-1999; 990US-0132048.

PR 30-APR-1999; 990US-0132407.

PR 04-MAY-1999; 990US-0132484.

PR 05-MAY-1999; 990US-0134485.

PR 06-MAY-1999; 990US-0132486.

PR 06-MAY-1999; 990US-0132487.

PR 07-MAY-1999; 990US-0132863.

PR 11-MAY-1999; 990US-0134256.

PR 990US-0134256.

PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137722.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	13-AUG-1999;	99US-0148565.
PR	10-JUN-1999;	99US-0138540.	PR	13-AUG-1999;	99US-0148684.
PR	10-JUN-1999;	99US-0138847.	PR	16-AUG-1999;	99US-0149735.
PR	14-JUN-1999;	99US-0139119.	PR	16-AUG-1999;	99US-0149735.
PR	16-JUN-1999;	99US-0139452.	PR	17-AUG-1999;	99US-0149717.
PR	16-JUN-1999;	99US-0139453.	PR	17-AUG-1999;	99US-0149717.
PR	17-JUN-1999;	99US-0139492.	PR	11-AUG-1999;	99US-0149319.
PR	18-JUN-1999;	99US-0139454.	PR	12-AUG-1999;	99US-0149341.
PR	18-JUN-1999;	99US-0139455.	PR	13-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139456.	PR	13-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139462.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139463.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139460.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139462.	PR	18-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139463.	PR	18-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139464.	PR	20-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139465.	PR	20-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139750.	PR	20-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139763.	PR	20-AUG-1999;	99US-0149368.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0149902.
PR	22-JUN-1999;	99US-0139899.	PR	27-AUG-1999;	99US-0149902.
PR	23-JUN-1999;	99US-0140353.	PR	23-AUG-1999;	99US-0149930.
PR	24-JUN-1999;	99US-0140354.	PR	25-AUG-1999;	99US-0150566.
PR	02-JUL-1999;	99US-0142055.	PR	07-SEP-1999;	99US-0153363.
PR	06-JUL-1999;	99US-0142390.	PR	10-SEP-1999;	99US-0155070.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155758.
PR	09-JUN-1999;	99US-014287.	PR	15-SEP-1999;	99US-0153303.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-015438.
PR	13-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154930.
PR	14-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0154779.
PR	15-JUL-1999;	99US-0143624.	PR	22-SEP-1999;	99US-0151339.
PR	16-JUL-1999;	99US-0144005.	PR	23-SEP-1999;	99US-015486.
PR	16-JUL-1999;	99US-0144085.	PR	24-SEP-1999;	99US-0156559.
PR	16-JUL-1999;	99US-0144290.	PR	28-SEP-1999;	99US-015458.
PR	19-JUL-1999;	99US-014297.	PR	29-SEP-1999;	99US-015039.
PR	19-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0151779.
PR	19-JUL-1999;	99US-0144332.	PR	05-OCT-1999;	99US-0157753.
PR	19-JUL-1999;	99US-0144333.	PR	06-OCT-1999;	99US-015865.
PR	19-JUL-1999;	99US-0144334.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144335.	PR	08-OCT-1999;	99US-0158232.
PR	20-JUL-1999;	99US-01443352.	PR	12-OCT-1999;	99US-0153369.
PR	20-JUL-1999;	99US-0144632.	PR	13-OCT-1999;	99US-0159293.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	13-OCT-1999;	99US-015295.
PR	21-JUL-1999;	99US-0145086.	PR	14-OCT-1999;	99US-0159329.
PR	21-JUL-1999;	99US-0145086.	PR	14-OCT-1999;	99US-0159330.
PR	22-JUL-1999;	99US-0145087.	PR	14-OCT-1999;	99US-015331.
PR	22-JUL-1999;	99US-0145087.	PR	27-JUL-1999;	99US-0145918.
PR	22-JUL-1999;	99US-0145089.	PR	27-JUL-1999;	99US-0145919.
PR	22-JUL-1999;	99US-0145192.	PR	27-JUL-1999;	99US-0145951.
PR	23-JUL-1999;	99US-0145145.	PR	02-AUG-1999;	99US-0146386.
PR	23-JUL-1999;	99US-0145218.	PR	02-AUG-1999;	99US-0146388.
PR	23-JUL-1999;	99US-0146389.	PR	02-AUG-1999;	99US-0146389.
PR	23-JUL-1999;	99US-0146767.	PR	26-JUL-1999;	99US-0146768.
PR	23-JUL-1999;	99US-0160770.	PR	27-JUL-1999;	99US-0160770.
PR	23-JUL-1999;	99US-0160814.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0160815.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0160981.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0160981.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0160981.	PR	22-OCT-1999;	99US-0160981.

	PR	22-OCT-1999;	99US -0160989.
	PR	25-OCT-1999;	99US -0161404.
	PR	25-OCT-1999;	99US -0161405.
	PR	26-OCT-1999;	99US -0161359.
	PR	26-OCT-1999;	99US -0161360.
	PR	28-OCT-1999;	99US -0161920.
	PR	28-OCT-1999;	99US -0161992.
	PR	28-OCT-1999;	99US -0161993.
	PR	29-OCT-1999;	99US -0162142.
Query Match		99.6%	Score 5654; DB 21; Length 1065;
Best Local Similarity		99.7%	Pred. No. 0; Mismatches 3; Indels 0; Gaps 0
Matches		1062;	Conservative 0; MisMatches 0; Indels 0; Gaps 0
Qy	1	MSESEGEGTAGKPMKNIVPOTCQICSDNVNGVTKVGDGDRFVACDICSFVCRCPEYERKGNO	60
Db	61	SCPQCKTRVKLGSPATPGDKDGLADEGTVENYPKEKISERMLGHILTRKGEM	120
	1	MSESEGEGTAGKPMKNIVPOTCQICSDNVNGVTKVGDGDRFVACDICSFVCRCPEYERKGNO	60
Qy	61	SCPQCKTRVKLGSPATPGDKDGLADEGTVENYPKEKISERMLGHILTRKGEM	120
Db	61	SCPQCKTRVKLGSPATPGDKDGLADEGTVENYPKEKISERMLGHILTRKGEM	120
Qy	121	GERQYDKEVSHNHLPLRTSQRDQSFGESASPERLVSSTAGGERLPLYSVDNSPNNR	180
Db	121	GERQYDKEVSHNHLPLRTSQRDQSFGESASPERLVSSTAGGERLPLYSVDNSPNNR	180
Qy	181	IYDPVGVLGNVAWERVDAWKMOKBKTGVSTOASERSGVIDASTDILADEALINDEA	240
Db	181	IYDPVGVLGNVAWERVDAWKMOKBKTGVSTOASERSGVIDASTDILADEALINDEA	240
Qy	241	RQLRSKYSIPSSRNPYRMVIMRLVILCFLYRITNPVNAFLWLYVICIWFAL	300
Db	241	RQLRSKYSIPSSRNPYRMVIMRLVILCFLYRITNPVNAFLWLYVICIWFAL	300
Qy	301	SWILDQFFKWPVNRETIDLRALRYDREGEPSOLAAVDIFVSTDPLKPLPVNTVL	360
Db	301	SWILDQFFKWPVNRETIDLRALRYDREGEPSOLAAVDIFVSTDPLKPLPVNTVL	360
Qy	361	STIAYDVPDKSCYVFDGAMMFSFESLAETSFARKWPKCKYSIERRAPENYFAAK	420
Db	361	STIAYDVPDKSCYVFDGAMMFSFESLAETSFARKWPKCKYSIERRAPENYFAAK	420
Qy	421	IDVLKDQVTSFKDRRAMKREYEEFKIRINALYSKALKCPEEGNMQDGTPWPNNTGD	480
Db	421	IDVLKDQVTSFKDRRAMKREYEEFKIRINALYSKALKCPEEGNMQDGTPWPNNTGD	480
Qy	481	HPOMIQVLGONGLDAEAGNELPRIWYREKRGFOHHKKAGANALVRSAVITNGPF	540
Db	481	HPOMIQVLGONGLDAEAGNELPRIWYREKRGFOHHKKAGANALVRSAVITNGPF	540
Qy	541	TLMDCDHYINNSKALRAMECMFLDNPUGKOVCTVOPPORFDGTDKNDYANRNFTEDI	600
Db	541	TLMDCDHYINNSKALRAMECMFLDNPUGKOVCTVOPPORFDGTDKNDYANRNFTEDI	600
Qy	601	NLRGLDGIOGPVYVGTCGVFNRTALYGEPPPIKVKHKKPSLLSKCGSRKKNSKAKES	660
Db	601	NLRGLDGIOGPVYVGTCGVFNRTALYGEPPPIKVKHKKPSLLSKCGSRKKNSKAKES	660
Qy	661	DKKKSGRHTDSTPVFNLDIEEGVEAGFDDELAKLMSLERSERFGQSAVFNSTIME	720
Db	661	DKKKSGRHTDSTPVFNLDIEEGVEAGFDDELAKLMSLERSERFGQSAVFNSTIME	720
Qy	721	NGCVPSATPENLIEKAHVISGYEDKCSDWGMEGWIGWGSYUTEILTGSKMHAGWRSI	780
Db	721	NGCVPSATPENLIEKAHVISGYEDKCSDWGMEGWIGWGSYUTEILTGSKMHAGWRSI	780
Qy	781	YCMPLPAFKGSPAPINSLDRNLQVRLWALGSVEILSRSRHOPIWCGYNGRKFLERAYN	840
Db	781	YCMPLPAFKGSPAPINSLDRNLQVRLWALGSVEILSRSRHOPIWCGYNGRKFLERAYN	840
Qy	841	TTTYPISIPLMCTLAVCLFTNQFLPOISNTASINTLSLEISIATGILEMRSGV	900

RESULT 3

ABB93356 ABB93356 standard; Protein; 1065 AA.

ID XX

AC ABB93356;

XX DT 31-MAY-2002 (first entry)

XX DE Herbicidally active polypeptide SEQ ID NO 2567.

XX KW Herbicidal; plant; agriculture; herbicide.

XX OS Arabidopsis thaliana.

XX PN WO200210210-A2.

XX PD 07-FEB-2002.

XX PF 28-AUG-2001; 2001WO-EP09892.

XX PR 28-AUG-2001; 2001WO-EP09892.

XX PA (FARB) BAYER AG.

XX PI Tietjen K, Weidler M;

DR WPI; 2002-269010/31.

XX

XX PT Identifying plant target proteins for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms -

XX PS Claim 5; SEQ ID NO 2567; 261PP + Sequence Listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016) for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides.

XX SQ Sequence 1065 AA;

Query Match 99.6%; Score 5654; DB 23; Length 1065;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1062; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 MESEGETAGPKPKNNIVQTCQICSDNKGKTVGDRFVACDCISFPVCPRCYERKDQGQ 60
1 MESEGETAGPKPKNNIVQTCQICSDNKGKTVGDRFVACDCISFPVCPRCYERKDQGQ 60

Db 61 SCPOCKKRYKJGKSPAIPGDKEDEGLADEGVENPQEKIERSNLGWHLTKGEEM 120

Db 61 SQPQCKKRYKRKGSPAPGDKDDEGLADEGVFENFFQKEKTSERMLGWHLTRGKEM 120 DT 18-OCT-2000 (first entry)
 Qy ||||||| 121 GEROYDKEVSHNHLPLRISLRODTSGEFAASPERLYSTIAGKRLPYSSDVNQSNRR 180 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 61574.
 Db 121 GEROYDKEVSHNHLPLRISLRODTSGEFAASPERLYSTIAGKRLPYSSDVNQSNRR 180 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 Qy ||||||| 181 IVDPGLGNNVAKERDVGQWKMQEKTNGPVSTQASERGGVVIDASDITLADELLNDEA 240 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 Db 181 IVDPGLGNNVAKERDVGQWKMQEKTNGPVSTQASERGGVVIDASDITLADELLNDEA 240 OS termination sequence.
 Qy ||||||| 241 ROPSLRKVISPPSRINPDRYRMVIMRLVILCFLHYRINTNPVNNAFAWLMVSICETWFAL 300 XX Arabidopsis thaliana.
 Db 241 ROPSLRKVISPPSRINPDRYRMVIMRLVILCFLHYRINTNPVNNAFAWLMVSICETWFAL 300 XX PR EPI033405-A2.
 Qy ||||||| 301 SWILDQFKWFPVNRERYLTLRALKYDREGEPEPSQLAAWDFVSTDPLKEPPLTAVNL 360 PD 06-SEP-2000.
 Db 301 SWILDQFKWFPVNRERYLTLRALKYDREGEPEPSQLAAWDFVSTDPLKEPPLTAVNL 360 XX PR 25-FEB-1999; 99US-0121825.
 Qy ||||||| 361 STLAQDPVFKVSCYVFDGAMLSFDSLAESEFARKWPCKKSYIEPRAPEWFAK 420 PR 09-MAR-1999; 99US-0123548.
 Db 361 STLAQDPVFKVSCYVSDGAAMLSFDSLAESEFARKWPCKKSYIEPRAPEWFAK 420 PR 23-MAR-1999; 99US-0123180.
 Qy ||||||| 421 IDYLDKDVQTSVKDRRAKREYEEFKTRINALVSKAKLCPEEGWMQDGTPWPGNNIGD 480 PR 05-MAR-1999; 99US-0122788.
 Db 421 IDYLDKDVQTSVKDRRAKREYEEFKTRINALVSKAKLCPEEGWMQDGTPWPGNNIGD 480 PR 25-MAR-1999; 99US-0122578.
 Qy ||||||| 541 TLNDQDHYINNSKALREAMCTLMDPQGKOCYVQPFQREGIDKNDRYARNRTVEDI 600 PR 29-MAR-1999; 99US-0126264.
 Db 541 TLNDQDHYINNSKALREAMCTLMDPQGKOCYVQPFQREGIDKNDRYARNRTVEDI 600 PR 01-APR-1999; 99US-0127462.
 Qy ||||||| 601 NURGLDQIQGPWYQGQCVFNITALQYEPPKVKHHKPSLISKLCGGSRKANSKAKES 660 PR 06-APR-1999; 99US-0128234.
 Db 601 NLRGLDQIQGPWYQGQCVFNITALQYEPPKVKHHKPSLISKLCGGSRKANSKAKES 660 PR 08-APR-1999; 99US-0128714.
 Qy ||||||| 661 DKKSGRHTDSTPVFNEDDIEFEGVEAGAFDDEKALLMSQMLEKRTGQSAVFASTIME 720 PR 16-APR-1999; 99US-0129845.
 Db 661 DKKSGRHTDSTPVFNEDDIEFEGVEAGAFDDEKALLMSQMLEKRTGQSAVFASTIME 720 PR 19-APR-1999; 99US-0130077.
 Qy ||||||| 721 NGGVPEPATPENILKEIAHVHSCGYEDKSDMGEIWTGYSYTEDILGKHKHARWSI 780 PR 21-APR-1999; 99US-0130449.
 Db 721 NGGVPEPATPENILKEIAHVHSCGYEDKSDMGEIWTGYSYTEDILGKHKHARWSI 780 PR 23-APR-1999; 99US-0130510.
 Qy ||||||| 781 YCMPLKUPAFKGSSAPINLSDRLQVLRWALGSVEILSRRHCPIWYGNGRLKLERAYN 840 PR 23-APR-1999; 99US-0130891.
 Db 781 YCMPLKUPAFKGSSAPINLSDRLQVLRWALGSVEILSRRHCPIWYGNGRLKLERAYN 840 PR 28-APR-1999; 99US-0131149.
 Qy ||||||| 841 TTIPITSIPLIMCYTLLAVCLIFTNQFIPOQINSIASIWFSLFLSTFATGLEMRWSGV 900 PR 30-APR-1999; 99US-0132048.
 Db 841 TTIPITSIPLIMCYTLLAVCLIFTNQFIPOQINSIASIWFSLFLSTFATGLEMRWSGV 900 PR 19-MAY-1999; 99US-0132407.
 Qy ||||||| 901 GIDEWWRNEQFWVIGVSAHLFAFGQCLIKVLAGIDNTFTVSKASDEDGPAELLFK 960 PR 21-MAY-1999; 99US-0132484.
 Db 901 GIDEWWRNEQFWVIGVSAHLFAFGQCLIKVLAGIDNTFTVSKASDEDGPAELLFK 960 PR 21-MAY-1999; 99US-0132485.
 Qy ||||||| 961 TTLLIPTTLLIUNLVGIVAGVSYAINSGYQSWGPLEGKLFCAFWMVWHLYPLKGLMR 1020 PR 23-MAY-1999; 99US-0132486.
 Db 961 TTLLIPTTLLIUNLVGIVAGVSYAINSGYQSWGPLEGKLFCAFWMVWHLYPLKGLMR 1020 PR 06-MAY-1999; 99US-0132487.
 Qy ||||||| 1021 ONRTPITIVVWMSVLLASIFSLWNRIDPFTSRVTGPDILECGNC 1065 PR 07-MAY-1999; 99US-0132863.
 Db 1021 ONRTPITIVVWMSVLLASIFSLWNRIDPFTSRVTGPDILECGNC 1065 PR 11-MAY-1999; 99US-0132864.
 Qy ||||||| 1021 ONRTPITIVVWMSVLLASIFSLWNRIDPFTSRVTGPDILECGNC 1065 PR 15-MAY-1999; 99US-0132865.
 Db 1021 ONRTPITIVVWMSVLLASIFSLWNRIDPFTSRVTGPDILECGNC 1065 PR 19-MAY-1999; 99US-0132866.
 RESULT 4 PR 18-OCT-2000 (first entry)
 AAG48734 PR 18-OCT-2000 (first entry)
 ID AAG48734 standard: Protein: 1054 AA. PR 18-OCT-2000 (first entry)
 AC AAG48734; PR 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 61574.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 XX hybridisation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 XX OS Arabidopsis thaliana.
 XX PR EPI033405-A2.
 XX PD 06-SEP-2000.
 XX PR 25-FEB-2000; 2000EP-0301439.
 XX PR 09-MAR-1999; 99US-0121825.
 XX PR 23-MAR-1999; 99US-0122578.
 XX PR 25-MAR-1999; 99US-0126264.
 XX PR 29-MAR-1999; 99US-0126785.
 XX PR 01-APR-1999; 99US-0127462.
 XX PR 06-APR-1999; 99US-0128234.
 XX PR 08-APR-1999; 99US-0128714.
 XX PR 16-APR-1999; 99US-0129845.
 XX PR 19-APR-1999; 99US-0130077.
 XX PR 21-APR-1999; 99US-0130449.
 XX PR 23-APR-1999; 99US-0130510.
 XX PR 23-APR-1999; 99US-0130891.
 XX PR 08-APR-1999; 99US-0131149.
 XX PR 16-APR-1999; 99US-0132048.
 XX PR 19-APR-1999; 99US-0132407.
 XX PR 21-APR-1999; 99US-0132484.
 XX PR 04-MAY-1999; 99US-0132485.
 XX PR 05-MAY-1999; 99US-0132486.
 XX PR 06-MAY-1999; 99US-0132487.
 XX PR 07-MAY-1999; 99US-0132863.
 XX PR 11-MAY-1999; 99US-0132864.
 XX PR 15-MAY-1999; 99US-0132865.
 XX PR 19-MAY-1999; 99US-0132866.
 XX PR 23-MAY-1999; 99US-0132867.
 XX PR 27-MAY-1999; 99US-0136392.
 XX PR 28-MAY-1999; 99US-0136782.
 XX PR 21-MAY-1999; 99US-0135353.
 XX PR 24-MAY-1999; 99US-0135629.
 XX PR 25-MAY-1999; 99US-0136021.
 XX PR 27-MAY-1999; 99US-0136392.
 XX PR 28-MAY-1999; 99US-0136782.
 XX PR 01-JUN-1999; 99US-0137222.
 XX PR 03-JUN-1999; 99US-0137528.
 XX PR 04-JUN-1999; 99US-0137529.
 XX PR 07-JUN-1999; 99US-0137724.
 XX PR 08-JUN-1999; 99US-0138094.
 XX PR 10-JUN-1999; 99US-0138540.
 XX PR 10-JUN-1999; 99US-0138847.
 XX PR 14-JUN-1999; 99US-0139119.
 XX PR 16-JUN-1999; 99US-0139452.
 XX PR 16-JUN-1999; 99US-0139453.
 XX PR 17-JUN-1999; 99US-0139492.
 XX PR 18-JUN-1999; 99US-0139454.
 XX PR 18-JUN-1999; 99US-0139455.
 XX PR 18-JUN-1999; 99US-0139456.
 XX PR 18-JUN-1999; 99US-0139457.
 XX PR 18-JUN-1999; 99US-0139458.
 XX PR 18-JUN-1999; 99US-0139459.
 XX PR 18-JUN-1999; 99US-0139460.
 XX PR 18-JUN-1999; 99US-0139461.
 XX PR 18-JUN-1999; 99US-0139462.
 XX PR 18-JUN-1999; 99US-0139463.
 XX PR 18-JUN-1999; 99US-0139750.

PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139893.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	24-JUN-1999;	99US-0140695.	PR	01-SEP-1999;	99US-0151930.
PR	28-JUN-1999;	99US-0140823.	PR	07-SEP-1999;	99US-0152363.
PR	29-JUN-1999;	99US-0140991.	PR	10-SEP-1999;	99US-0153070.
PR	30-JUN-1999;	99US-0141287.	PR	13-SEP-1999;	99US-0153758.
PR	01-JUL-1999;	99US-0141842.	PR	15-SEP-1999;	99US-0154018.
PR	02-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	06-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	08-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	09-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155486.
PR	12-JUL-1999;	99US-0142977.	PR	24-SEP-1999;	99US-0156559.
PR	13-JUL-1999;	99US-0143542.	PR	28-SEP-1999;	99US-0156458.
PR	14-JUL-1999;	99US-0143624.	PR	29-SEP-1999;	99US-0156596.
PR	15-JUL-1999;	99US-0144005.	PR	04-OCT-1999;	99US-0157117.
PR	16-JUL-1999;	99US-0144085.	PR	05-OCT-1999;	99US-0157753.
PR	17-JUL-1999;	99US-0144290.	PR	06-OCT-1999;	99US-0157865.
PR	19-JUL-1999;	99US-0144325.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144352.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144334.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144505.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144532.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144684.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159295.
PR	21-JUL-1999;	99US-0145086.	PR	19-OCT-1999;	99US-0159329.
PR	21-JUL-1999;	99US-0145089.	PR	20-OCT-1999;	99US-0159330.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0159331.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0159637.
PR	22-JUL-1999;	99US-0145195.	PR	21-OCT-1999;	99US-0159638.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0159684.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-01606989.
PR	22-JUL-1999;	99US-0145226.	PR	25-OCT-1999;	99US-0161040.
PR	22-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	21-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	22-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147260.	Query Match	98.6%	Score 5597; DB 21; Length 1054;
PR	06-AUG-1999;	99US-0147303.	Best Local Similarity	99.7%	Pred. No. 0; Mismatches 0; Gaps 0;
PR	06-AUG-1999;	99US-0147416.	Matches 1051; Conservative	0;	Mismatches 3; Index 0;
PR	09-AUG-1999;	99US-0147493.	Matches 1051;		
PR	10-AUG-1999;	99US-0148171.	Qy	12	MKNIVPOTCQICSDNNGKUWGDPSRFVACDICSCFPCVCPCEYERKDENOSCPCKTRYKR 71
PR	11-AUG-1999;	99US-0148319.	Db	1	MKNIVPOTCQICSDNNGKUWGDPSRFVACDICSCFPCVCPCEYERKDENOSCPCKTRYKR 60.
PR	12-AUG-1999;	99US-0148341.	Qy	72	LKSPPAIPCDKEDGDLADEGTVFENPYKEKISERMIGWHLTRGKGEMGEPOYDKVSH 131
PR	13-AUG-1999;	99US-0148365.	Db	61	LKSPPAIPCDKEDGDLADEGTVFENPYKEKISERMIGWHLTRGKGEMGEPOYDKVSH 120
PR	13-AUG-1999;	99US-0148684.	Qy	132	NHJPLRTRQDTSGEFSASPERLVSFTIAGGRKLYSSDNQNSPNRNRIVDPVGJNA 191
PR	16-AUG-1999;	99US-0149368.	Db	181	WKEVDGKMKQEKNTGKVSTOAESRGVTDASTDILADERLNDNEAROPLRSKYSIP 240
PR	17-AUG-1999;	99US-0149375.	Qy	252	SSINPYKWMIRLWICLFLHYRINPVNPNAFLMUVSCEIWALSMWLDQPKWF 311
PR	18-AUG-1999;	99US-0149426.	Db		
PR	20-AUG-1999;	99US-0149722.	Qy	192	WKEVDGKMKQEKNTGKVSTOAESRGVTDASTDILADERLNDNEAROPLRSKYSIP 251
PR	23-AUG-1999;	99US-0149902.	Db		
PR	23-AUG-1999;	99US-0149930.	Qy	181	WKEVDGKMKQEKNTGKVSTOAESRGVTDASTDILADERLNDNEAROPLRSKYSIP 240
PR	25-AUG-1999;	99US-0150566.	Db		
PR	26-AUG-1999;	99US-0150884.	Qy		
PR	27-AUG-1999;	99US-0151085.	Db		

Db 241 SSRINPYRMVIMRLVILCFLFHRYTNPVNAFAILWLVSVICEIFALPSNILDQFPKW 300 PD 06-SEP-2000.
 Qy |||||||PVNRETYLDRIALRYDREGESQLAAVDIFYSTVDLKEPPLVTAVLSLAVYPVK 371 XX
 Db 301 PVNRETYLDRIALRYDREGESQLAAVDIFYSTVDLKEPPLVTAVLSLAVYPVK 360 PR 25-FEB-2000; 2000EP-0301439.
 Qy 372 VSCIVYDDGAAMLFESLAETSEFARKWVPPCKKSISIEPRAPEWTPAKIDYLKDQVTS 431 PR 05-MAR-1999; 99US-0123180.
 Db 361 VSCIVYDDGAAMLFESLAETSEFARKWVPPCKKSISIEPRAPEWTPAKIDYLKDQVTS 420 PR 09-MAR-1999; 99US-0123185.
 Qy 432 FVKDRRRAKREVEEKFIRTRNALVSKALKCPERGWWMQDGTPWPGNNNGDHGMIOVFLGQ 491 PR 23-MAR-1999; 99US-0123188.
 Db 421 FVKDRRRAKREVEEKFIRTRNALVSKALKCPERGWWMQDGTPWPGNNNTRDHPGMIOVFLGQ 480 PR 25-MAR-1999; 99US-0123192.
 Qy 492 NGGLDEEGNELPRLVYVSREKRGCFCHHKAGAMNALVYRSAVINGPFTNLDCDHY IN 551 PR 29-MAR-1999; 99US-0123194.
 Db 481 NGGLDEEGNELPRLVYVSREKRGCFCHHKAGAMNALVYRSAVINGPFTNLDCDHY IN 540 PR 01-APR-1999; 99US-0123196.
 Qy 552 NSKLREAMCFLMDPNIKGQCVYQFQRCIDKNDRYYANRNTFFDINURGLDQIGP 611 PR 05-APR-1999; 99US-0123199.
 Qy |||||||NSKLREAMCFLMDPNIKGQCVYQFQRCIDKNDRYYANRNTFFDINURGLDQIGP 600 PR 08-APR-1999; 99US-0123204.
 Db 541 NSKLREAMCFLMDPNIKGQCVYQFQRCIDKNDRYYANRNTFFDINURGLDQIGP 600 PR 19-APR-1999; 99US-0123208.
 Qy 612 WVGTCGVNRITALYGEPEPTIKVHKRPLSLSKLCGGSRKNSKAKKESDKKKSGRTS 671 PR 21-APR-1999; 99US-0130049.
 Db 601 WVGTCGVFNKITALYGEPEPTIKVHKRPLSLSKLCGGSRKNSKAKKESDKKKSGRTS 660 PR 23-APR-1999; 99US-0130310.
 Qy 672 TVPVFENLDIEBEGVEGAGFEDDEKALLMSOMSLERFGQSAVFASLTENGVPPSATPE 731 PR 28-APR-1999; 99US-0130891.
 Db 661 TVPVFENLDIEBEGVEGAGFEDDEKALLMSOMSLERFGQSAVFASLTENGVPPSATPE 720 PR 30-APR-1999; 99US-0130895.
 Qy 732 NLUKEHTHVISCGYEDKSDMGMEIGHTYGSYTEDITGKMHARGRSTICMPLKAFCG 791 PR 19-APR-1999; 99US-0132048.
 Db 721 NILKEA1HVTS CGYEDKSDMGMEIGHTYGSYTEDITGKMHARGRSTICMPLKAFCG 780 PR 21-APR-1999; 99US-0132056.
 Qy 792 SAPINUSDRNLNOVLRVALGSWEILFSRRCPTWGYGNRKKLERFERAVNTIYTPIPL 851 PR 04-MAY-1999; 99US-0132484.
 Db 781 SAPINUSDRNLNOVLRVALGSWEILFSRRCPTWGYGNRKKLERFERAVNTIYTPIPL 840 PR 05-MAY-1999; 99US-0132485.
 Qy 852 LMYCTLLAVCLPNTQFQIPOINSIASIWIFLSFLSIFATGILEMRWMSGVGE 911 PR 06-MAY-1999; 99US-0132486.
 Db 841 LMYCTLLAVCLPNTQFQIPOINSIASIWIFLSFLSIFATGILEMRWMSGVGE 900 PR 07-MAY-1999; 99US-0132487.
 Qy 912 WVGGSYAHFLAVFOGQILKVLLAGIDNFNTVSKASDEDGDAELYFKWTLLIPTLL 971 PR 07-MAY-1999; 99US-0134449.
 Db 901 WVGGSYAHFLAVFOGQILKVLLAGIDNFNTVSKASDEDGDAELYFKWTLLIPTLL 960 PR 08-MAY-1999; 99US-0134450.
 Qy 972 IVNLVGVWAGSYAINGSYQSWGPLGKLFPAFWVVLHLYFLKGJMGQRNRTPTVVW 1031 PR 01-JUN-1999; 99US-0134451.
 Db 961 IVNLVGVWAGSYAINGSYQSWGPLGKLFPAFWVVLHLYFLKGJMGQRNRTPTVVW 1020 PR 02-JUN-1999; 99US-0134452.
 Qy 1032 SYLLASIFSLLWVRIDPFTSRVTGPDPILLECCINC 1065 PR 03-JUN-1999; 99US-0134453.
 Db 1021 SYLLASIFSLLWVRIDPFTSRVTGPDPILLECCINC 1054 PR 04-JUN-1999; 99US-0134454.
 RESULT 5 PR 05-JUN-1999; 99US-0134455.
 AAG48735 ID AAG48735 standard; Protein; 959 AA. PR 06-JUN-1999; 99US-0134456.
 XX AC AAG48735; PR 07-JUN-1999; 99US-0134457.
 XX DT 18-OCT-2000 (first entry) PR 08-JUN-1999; 99US-0134458.
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 61575. PR 09-JUN-1999; 99US-0134459.
 KW Protein identification; signal transduction pathway; metabolic pathway; PR 10-JUN-1999; 99US-0134460.
 KW hybridisation assay; genetic mapping; gene expression control; promoter; PR 11-JUN-1999; 99US-0134461.
 KW termination sequence. PR 12-JUN-1999; 99US-0134462.
 OS Arabidopsis thaliana. PR 13-JUN-1999; 99US-0134463.
 PN EP1033405-A2. PR 14-JUN-1999; 99US-0134463.
 XX PR 15-JUN-1999; 99US-0134463.
 PR 16-JUN-1999; 99US-0134463.
 PR 17-JUN-1999; 99US-0134463.
 PR 18-JUN-1999; 99US-0134463.
 PR 19-JUN-1999; 99US-0134463.
 PR 20-JUN-1999; 99US-0134463.
 PR 21-JUN-1999; 99US-0134463.
 PR 22-JUN-1999; 99US-0134463.
 PR 23-JUN-1999; 99US-0134463.
 PR 24-JUN-1999; 99US-0140554.
 PR 25-JUN-1999; 99US-0140554.
 PR 26-JUN-1999; 99US-0140554.
 PR 27-JUN-1999; 99US-0140554.
 PR 28-JUN-1999; 99US-0140554.
 PR 29-JUN-1999; 99US-0140554.
 PR 30-JUN-1999; 99US-0140554.
 PR 01-JUL-1999; 99US-0140554.
 PR 02-JUL-1999; 99US-0140554.
 XX PR 03-JUL-1999; 99US-0140554.

QY	587	NDRYARNRTTYFDINRLGLDGIQPVVVGTCVFNRTALYGEPPKVKHKPSLSQLC	480
CC			
Db	481	NDRYARNRTTYFDINRLGLDGIQPVVVGTCVFNRTALYGEPPKVKHKPSLSQLC	540
Db	541	GGSRKNSKAKKESDKKKSGRINTDSTPWFVNLDIIEEGVEAGAFDEVALMSQMSLER	706
Qy	541		
Db	707	FQGOSAVFVASTLMEENSPVPSATPENNLIKEATHVISCGYEDKSDWGMIGWIGSVTEDI	756
Qy	707		
Db	601	FQGOSAVFVASTLMEENSPVPSATPENNLIKEATHVISCGYEDKSDWGMIGWIGSVTEDI	660
Qy	601		
Db	767	LTGFKMHARGRSIYCMPLPAFKGSAPINSLDRNQVLWRALGSEVELFSRHCPIWQY	826
Qy	767		
Db	661	LTGFKMHARGRSIYCMPLPAFKGSAPINSLDRNQVLWRALGSEVELFSRHCPIWQY	720
Qy	827	NGRLKFLERFAYVNTTLYPITSIPLMLCTILAVCLFTNQFLFOISNTASINFLSFLS	886
Qy	827		
Db	721	NGRLKFLERFAYVNTTLYPITSIPLMLCTILAVCLFTNQFLFOISNTASINFLSFLS	780
Qy	887	IFATGILEMRNSGVGIDEMWWRHEQFWYIGGSRHLEAFQGILKVLAGIDTNFTVSAS	946
Qy	887		
Db	781	IFATGILEMRNSGVGIDEMWWRHEQFWYIGGSRHLEAFQGILKVLAGIDTNFTVSAS	840
Qy	947	DEDGDFEELYLELKWTMILLIPPTLLINLVWVAGVSAINSQWSMPLFGKLFFAFW	1006
Db	841	DEDGDFEELYLELKWTMILLIPPTLLINLVWVAGVSAINSQWSMPLFGKLFFAFW	900
Qy	1007	IYHLYPFKLGLMGRQRNRPTIVVWWSVILASIFSLWLWIRDPETSRVNGPDILECGINC	1065
Db	901	IYHLYPFKLGLMGRQRNRPTIVVWWSVILASIFSLWLWIRDPETSRVNGPDILECGINC	959
RESULT 5			
ID	AY84108	AAY84108 standard; Protein; 1077 AA.	
XX			
AC			
XX			
DT	03-JUL-2000	(first entry)	
XX			
DE	Amino acid sequence of a maize cellulose synthase.		
XX			
KW	Maize; cellulose synthase; stalk quality; strand; silage; cellulose; transgenic plant; plant breeding marker.		
XX			
OS	zea mays.		
XX			
Key	Location/Qualifiers		
FH	Misc-difference 494	/note= "encoded by MGG"	
FT	WO200009706-A2.		
FT	24-FEB-2000.		
FT	16-AUG-1999;	99WO-US18760.	
PR	17-AUG-1998;	98US-0096822.	
XX			
PA	(PION-) PIONEER HI-BRED INT INC.		
XX			
PI	Dhugga KS, Heentjars TG, Bowen BA, Wang X;		
XX			
WPI;	2000-224343/19.		
DR	N-PSBB; AAZ99434.		
XX			
New genes which encode maize cellulose synthase polypeptides in plants useful for modulating the expression of cellulose synthase in plants and to produce transgenic plants expressing the novel protein			
XX			
QY	582	DCIKNDRYARNRTTYFDINRLGLDGIQPVVVGTCVFNRTALYGEPPKVKHKPSL	641
Qy	597	DCIDRNDRYARNRTTYFDINRLGLDGIQPVVVGTCVFNRTALYGEPEPPI-KOKGG	654
DB	537	AGAMNALVRVAVLTINGPFIINLDCDIYINNSKALRERAMCFIAMDPNLGKQCVWQFFRF	591
DB	537		
DB	537	AGAMNALVRVAVLTINGPFIINLDCDIYINNSKALRERAMCFIAMDPNLGKQCVWQFFRF	591
DB	462	ERGWVMODGTPMPGNNGDHQPMIQVFLGQNGLDAEGLNELLRLVYRSREKRGFQHKK	521
DB	477	ERGWIMODGTPMPGNNGDHQPMIQVFLGQNGLDAEGLNELLRLVYRSREKRGFQHKK	536
Qy	522	AGAMNALVRVAVLTINGPFIINLDCDIYINNSKALRERAMCFIAMDPNLGKQCVWQFFRF	581
Qy	522		
DB	537	AGAMNALVRVAVLTINGPFIINLDCDIYINNSKALRERAMCFIAMDPNLGKQCVWQFFRF	591
DB	537		

QY 642 LSKLGGSGRSRKNSKAKEDDKKKSGRHTDSTVPDFNLDDIEEGVEAGFDDERKALLMSQM 701
 XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 CC The present sequence represents a maize cellulose synthase polypeptide.
 Db 655 LSSLCGG-RKKASKSKKGSKKSKSQKHDDSSVPVFNLEDIEEGVEAGFDDERKLLMSQM 713
 CC The cellulose synthase can be used for the improvement of stalk quality
 QY 702 SLEERFGGOSRAFVASTLMEENGGVPPSATEPENLKEAIIHVTCSCYEDKEDWGMIGWIGS 761
 CC for improved stand or silage. It also provides an increased concentration
 Db 714 SLERFRGGOSRAFVASTLMEENGGVPPSATEPENLKEAIIHVTCSCYEDKEDWGMIGWIGS 773
 CC of cellulose in the pericarp, hardening the kernel and improving its
 handling ability. The sequences are used to produce transgenic plants
 CC and seeds expressing the cellulose synthase. The polynucleotide is
 used for modulating, preferably increasing, the level of the synthase
 in a plant cell. The plants are preferably monocots. The polynucleotide
 is also used as a probe or primer in the detection quantitation or
 isolation of gene transcripts. The probes are useful in detecting
 Db 762 VTEDLTGKMHARGWRSYSCMPKLAPEGSAPINLSPRLQUNVRWAISGVELFSRHP 821
 CC deficiencies in the level of mRNA in screenings for desired transgenic
 CC plant, for detecting mutations in the gene, for monitoring upregulation
 CC of expression or changes in enzyme activity in screening assays of
 compounds, for detection of any number of allelic variants of the gene,
 Db 834 LWYGGYGRKFLEFAYNTVTPLSLPLVCLPAACLGKFIPPEISFASTWEI 893
 CC or for use as molecular markers in plant breeding programs. The polynucleotide
 QY 882 SLFLSIFATGILEMRHWSGIGDEWMRNEQFWITGGVSAILFLAVFOGILTILVLAGIDTNFTV 941
 CC isolated nucleic acids of the present invention can also be used for
 recombinant expression of their encoded polypeptides or for use as
 Db 894 SLFSIFATGILEMRHWSGIGDEWMRNEQFWITGGVSAILFLAVFOGILTILVLAGIDTNFTV 953
 CC immunogens in the preparation and/or screening of antibodies. The
 proteins can be employed in assays for enzyme agonists or antagonists
 QY 942 TSKASDEDDFDAEYLKWTLLIPTTLLIVLNULVGWWAGVSTAINGSKWPLGFLF 1001
 CC or for use of immunogens or antigens to obtain
 Db 955 TSASDEDDFDAEYLKWTLLIPTTLLIVLNULVGWWAGVSTAINGSKWPLGFLF 1013
 CC antibodies specifically immunoreactive with a protein.

XX Sequence 1077 AA:
 Query Match 80.5%; Score 4572; DB 21; Length 1077;
 Best Local Similarity 79.4%; Pred. No. 0;
 Matches 861; Conservative 81; Mismatches 112; Indels 30; gaps 10;
 XX
 Db 1062 GINC 1065
 1074 GINC 1077

RESULT 7

ID AAY84114 standard; Protein: 1077 AA.

XX
 AC AAY84114;
 XX
 DT 03-JUL-2000 (first entry)

XX Amino acid sequence of a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 XX transgenic plant; plant breeding marker.

OS zea mays.

XX
 Key Location/Qualifiers
 FT Misc-difference 494
 FT /note= "encoded by MGCG"

XX
 WO200005706-A2.

XX
 PD 24-FEB-2000.

XX
 PF 16-AUG-1999; 99WO-US18760.

PR 17-AUG-1998; 98US-0096822.

XX
 PA (PTON-) PIONEER HI-BRED INT INC.

XX
 PT Dugga RS, Helentjaris TG, Bowen BA, Wang X;

DR WPI; 2000-224343/-19.

XX N-PSDB; AAZ9512.

XX
 PT New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants
 PT and to produce transgenic plants expressing the novel protein

XX
 Claim 15; Page 142-144; 119pp; English.

Query Match 80.5%; Score 4572; DB 21; Length 1077;
 Best Local Similarity 79.4%; Pred. No. 0;
 Matches 861; Conservative 81; Mismatches 112; Indels 30; gaps 10;
 XX
 Db 117 GEEMGEPQYDK-----EVSNHLPLRTSRQDTSGEFSAASPERLSVSSTAGK 165
 119 SGDVGGRPKYDSGETGLTKYDSDGEIPRGYPTSVINSQ_ISGEIPGAQSPDHMMSPGNIGK 177

Db 166 RLPHYSSDVQSPNPRRRIVDPVGLGNWAWKERVQDVOKMKOEKNQCPV--STQASERGV- 221
 178 RAFFPY-VNHSPNPSPREPGSGSIGNAVKERVQDVOKMKDGKTLPMNTGTSAPSERGVG 236

Db 222 DIDASTDILADEALINDEARQPLSRKVSTPSSRINPYRNIMRLVILCFLHYRTPV 281
 237 DIDASTDINMEDALLNDETROPLSLRKVPLSSIINPYRNIVLRLVISFLHYRTPV 296

Db 282 PNAFAFLWIVSVCIEIWFAWSWILDQQPKWFPVNRETYLDLALRYDREGERSQALAVDIF 341
 297 RNAYPLWLWLSVICBIEIWFAWSWILDQQPKWFPNIRETYLDLALRYDREGERSQALAVDIF 356

Db 342 VSTVDPKLKEPPLVTANTVSLIAVDPYDVKVSYYFDGAAMLSFESEAETSEFARKWMP 401
 357 VSTVDPKLKEPPLVTANTVSLIAVDPYDVKVSYYFDGAAMLSFESEAETSEFARKWMP 416

Db 402 FCKYKSIERAPENYFAAKIDYJDKVOTSFVKDRKREEEFKIRINALVSKALKCP 461
 417 FVKKYKIERAPENYFSOKIDYJDKVAPSFVKDRKREEEFKIRVALVAKQKVP 476

Db 462 EEGWMODGTPWPGNNCDHPGMTOVFLGONGSLDAGSNELPRLYVVSREKRPGFHKK 521
 477 EEGWMODGTPWPGNTXDHPGMTOVFLGHSGGLDTEGNELPLRVVYSREKRPGFHKK 536

Db 522 AGAMNALVYSAVINGPILNLDCHYINNSKALREAMCFIMDPNGKQCYVQFQRF 581
 537 AGANNAFLRVYSAVINGPILNLDCHYINNSKALREAMCFIMDPNGKQCYVQFQRF 596

Db 582 DGIDKNDRYANRNFVFDTNLRGLGLOGPVYVGTGVCFNRALYGVPEPIVKHKPSL 641

Db 597 DGIDKNDRYANRNFVFDINLRGLGLOGPVYVGTGVCFNRALYGVPEPI-KQKGGE 654

Also disclosed are a recombinant expression cassette comprising the polynucleotide (operably linked to a promoter) a host cell comprising the recombinant expression cassette and a transgenic plant comprising the recombinant expression cassette. The nucleic acid is useful in the improvement of stalk quality for improved stand or silage, and in the increased concentration of cellulose in the pericarp, hardening the kernel, and thus improving its handling ability. The nucleic acids may also be used as probes or amplification primers in the detection, quantification or isolation of gene transcripts, as probes in detecting deficiencies in the level of mRNA, for detecting gene mutations or allelic variants, for monitoring up regulation of expression or changes in enzyme activity in screening assays, for site directed mutagenesis, and in sense or antisense suppression of one or more genes in a host cell, tissue or plant. The polypeptides may be used in assays for enzyme agonists or antagonists, and as immunogen or antigen to obtain antibodies specifically immunoreactive with the protein. The present sequence represents a corn cellulose synthase of the invention.

QY	642	LSKLGGSRKNSKAKKEDKKSSRQHDTSTVPPFNLDDETEGVGEGFDDDEKALIMSM
Db	655	LSSIGG-RKKASRSKKGSDKKSSQKHHVSSVPPVNELEDEEGVGAFGDFDEKSLMSQ
QY	702	SLEKRGOSAVESTLMNGGVPSATPENLKKIARHVSGYEDKSDMGMEIHYGS
Db	714	SLKRRFGOSRAFTSTLMEYGGVQSOATSEPSLKLKARHVSGYEDKTEWTEIGWIGS
QY	762	VTDILTGFKMHARGWSRICMPKPAFKGSAPINLSDRLNQVLRALGSEVILSRHCP
Db	774	VTDILTGFKMHARGWSRICMPKPAFKGSAPINLSDRLNQVLRALGSEVILSRHCP
QY	822	IWGYNGRILKFLERFAVNTIYTITSIPLIMCYCTTAVCLFTNQFIPQTSNIASTWFL
QY	834	LWGYGGALKFLERFAVNTIYTITSIPLIMCYCTTAVCLFTNQFIPQTSNIASTWFL
Db	882	SFLSISFATGILEMRHWGSGIDEMWRNEDFWVIGSYSAHLFAVEGFLKLVAGDINFTV
QY	894	SLTISIFATGILEMRHWGSGIDEMWRNEDFWVIGSYSAHLFAVEGFLKLVAGDINFTV
QY	942	TSKASDEGDFAELYLKWMUTLIPPTLLIVNLVWAGVSAYANSGYQSWGPLKLF
Db	954	TSKASDEGDFAELYLKWMUTLIPPTLLIVNLVWAGVSAYANSGYQSWGPLKLF
QY	1002	FATWWIVHLYPELKGLMGGRNRTPTIVWWVSVLIAISFLSLWVRLDPFTSRVTGPDLIC
Db	1014	FATWWIVHLYPELKGLMGGRNRTPTIVWWVSVLIAISFLSLWVRLDPFTSRVTGPDLIC
Oy	1062	GINC 1065
Db	1074	GINC 1077
RESULT 8		
ID	AAU10496	AAU10496 standard; Protein; 1079 AA.
XX	AC	AAU10496;
XX	DT	14-FEB-2002 (first entry)
DE	Corn cellulose synthase Cqrαel9/cesa-9.	
XX	KW	Corn; cellulose synthase; cdgs45; cesa-3; Cqrαel9; cesa-9; stalk quality; improved stand; silage; pericarp; kernel hardening; handling ability; transgenic plant; immunogen.
OS	zea mays.	
XX	PN	WO20179516-A2.
XX	PD	25-OCT-2001.
XX	PF	12-APR-2001; 2001WO-US11951.
PR		14-APR-2000; 2000US-0350483.
XX	PA	(PION-) PIONEER HI-BRED INT INC.
PI		Dhuga KS, Helentjaris TG;
XX	DR	WPI; 2002-0411338/05.
DR		N-PSDB; AS16458.
XX	PT	New cellulose synthase polypeptides and polynucleotides, useful in improving stalk quality or silage, and in increasing concentration of cellulose in the pericarp, hardening the kernel for improved handling ability -
PT	XX	Claim 3; Page 85-87; 88pp; English.
PS	XX	The invention relates to isolated nucleic acids encoding two cellulose synthase proteins from corn, Cdgs45 (cesa-3) and Cqrαel9 (cesa-9).

Also disclosed are a recombinant expression cassette comprising the polynucleotide (operably linked to a promoter) a host cell comprising the recombinant expression cassette and a transgenic plant comprising the recombinant expression cassette. The nucleic acid is useful in the improvement of stalk quality for improved stand or silage, and in the increased concentration of cellulose in the pericarp, hardening the kernel, and thus improving its handling ability. The nucleic acids may also be used as probes or amplification primers in the detection, quantification or isolation of gene transcripts, as probes in detecting deficiencies in the level of mRNA, for detecting gene mutations or allelic variants, for monitoring up regulation of expression or changes in enzyme activity in screening assays, for site directed mutagenesis, and in sense or antisense suppression of one or more genes in a host cell, tissue or plant. The polypeptides may be used in assays for enzyme agonists or antagonists, and as immunogen or antigen to obtain antibodies specifically immunoreactive with the protein. The present sequence represents a corn cellulose synthase of the invention.

Qy	827	NGRKLKLERFAVNTTIPISIPLMCTILAVCLFTNQFIPQISNIASTIWLSFLS	886	CC isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.
Db	836	NGRKLKLERFAVNTTIPISIPLMCTILAVCLFTNQFIPQISNIASTIWLSFLS	895	CC
Qy	887	IFATGILMRNSVGIDDEWWNEQFWNTTIPISIPLMCTILAVCLFTNQFIPQISNIASTIWLSFLS	946	CC
Db	896	IFATGILMRNSVGIDDEWWNEQFWNTTIPISIPLMCTILAVCLFTNQFIPQISNIASTIWLSFLS	955	CC
Qy	947	DEDGDAELYLIFKWTLILLIPPTLILVNLGVAGYSAHLLAVFOGLIKVLAGIDNFVSKAS	1006	XX
Db	956	DGDGAELYLIFKWTLILLIPPTLILVNLGVAGYSAHLLAVFOGLIKVLAGIDNFVSKAS	1015	SQ
RESULT	12			
AY84115				
ID	AY84115	standard; Protein; 1075 AA.		
XX				
AC	AY84115;			
XX				
DT	03-JUL-2000	(first entry)		
XX				
DE	Amino acid sequence of a maize cellulose synthase.			
XX				
KW	Maize; cellulose synthase; stalk quality; stand; silage; cellulose; transgenic plant; plant breeding marker.			
KW				
OS	zea mays.			
XX				
FH				
Key	Location/Qualifiers			
FT	Misc-difference 245	/note= "encoded by NAT"		
FT				
XX	WO200009706-A2.			
XX				
PD	24-FEB-2000.			
XX				
PF	16-AUG-1999;	99WO18760.		
PR	17-AUG-1998;	98US-009622.		
XX				
PA	(PION-) PIONEER HI-BRED INT INC.			
XX				
PI	Dhugga KS, Helentjaris TG, Bowen BA, Wang X;			
XX				
DR	WPI; 2000-224343/19.			
DR	N-PSDB; AN299515.			
PT	New genes which encode maize cellulose synthase polypeptides in plants and to produce transgenic plants expressing the novel protein			
PT	and to modulate the expression of cellulose synthase in plants			
PT	and to produce transgenic plants expressing the novel protein			
PS	Claim 15; Page 149-152; 119PP; English.			
The present sequence represents a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting gene deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The				
Qy	527	ALVRVSAVLNGPFILNUCDHYINNSKALREAMCFMLDPNLKGQCVYQFFRGDK	586	CC isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.
Db	539	ALTRVSAILNGAFLNVDCYFNNSKALREAMCFMLDPNLKGQCVYQFFRGDK	418	CC
Qy	407	SIERAPAEWYFAAKKIDYLKDQYQTSFVYKDRRAKMKREYEEFKIRINNALVSKALKCPEEGW	466	CC
Qy	419	NIERAPAEWYFAOKIDYLKDQYQTSFVYKERRAKMKREYEEFKIRINNALVSKALKCPEEGW	478	CC
Db	467	MQDGTPWPGNNYGDHPGMIQVQLGQNGQGLDAEGNELLRLRVYSSERERKPGFOHKKAGMN	526	CC
Qy	479	MADGTAWGNRPRDHPGMIQVLTGHSGLDQDGNELLRVYSSERERKPGFOHKKAGMN	538	CC
Db	527	ALVRVSAVLNGPFILNUCDHYINNSKALREAMCFMLDPNLKGQCVYQFFRGDK	586	CC
Qy	539	ALTRVSAILNGAFLNVDCYFNNSKALREAMCFMLDPNLKGQCVYQFFRGDK	418	CC
Db	587	NDRYANRNTVVEDIDNLRGDQIGGPVWVGTGCCVFRNTALYGEPEPKVKHKRPSLISKLC	646	CC
Qy	599	HDRYANRNTVVEDIDNLRGDQIGGPVWVGTGCCVFRNTALYGEPEPKVKHKRPSLISKLC	658	CC
Qy	647	GSSRKISKAKKEKKSGRTDSTPVFNDDIEGEGAFDDERKALMSMSLEK	706	CC
Db	659	CERRKKNSYMDQSQRIMKR-TESSARIFHNEDIEGEGE-GYEDERSVLMQSKLKE	715	CC
Qy	707	FGOSAVAVSTLMEENGVPPSATPENLKEATHVSGYEDKSDWMEIGWIGSYAEDI	766	CC
Db	716	FGOSPIFASTMTQGIPSPNPAJSIKEATHVSGYEDKTEWGEIGWIGSYAEDI	775	CC
Qy	767	LTCFKMHARGWISYCHPKLPARKGSAPINLUSDRNLQYRNLWLGSVLFSRHCPIHY	826	CC
Db	776	LTCFKMHARGWISYCHPKLPARKGSAPINLUSDRNLQYRNLWLGSVLFSRHCPIHY	835	CC
Qy	827	NERLKLFERFAVNTTIPISIPLMCTILAVCLFTNQFIPQISNIASTIWLSFLS	886	CC
Db	836	NGRKLKLERFAVNTTIPISIPLMCTILAVCLFTNQFIPQISNIASTIWLSFLS	895	CC

Db 956 DEGDFEALYVFRWTSLLIPPTVLVNLGVAGISAYNSQGQSWGPRLCKLFFSIWV 1015
 Qy 1007 IVHYPFLKGLMGRQNRPPTIVVWSVLLASIFSLWWRDFTSRVTPGDIL-EGCINC 1065
 Db 1016 IHLYPFLKGLMGRONRPTIVVWSVLLASIFSLWVKIDFISPDKAAALGQCGVNC 1075

RESULT 14
 AAY84110
 ID AAY84110 standard; protein: 1074 AA.
 AC AAY84110;
 XX DT 03-JUL-2000 (first entry)
 DE Amino acid sequence of a maize cellulose synthase.
 KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 KW transgenic plant; plant breeding marker.
 OS zea mays.
 XX PN WO200009705-A2.
 XX PD 24-FEB-2000.
 XX PF 16-AUG-1999; 99WO-US18760.
 PR 17-AUG-1998; 98US-0096822.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
 XX DR WPI; 2000-224343/19.
 DR N-PSDB; AA299500.
 XX PT New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants
 PT and to produce transgenic plants expressing the novel protein -
 XX PS Claim 15; Page 110-113; 119pp; English.

The present sequence represents a maize cellulose synthase polypeptide.
 CC The cellulose synthase can be used for the improvement of stalk quality
 CC for improved stand or silage. It also provides an increased concentration
 CC of cellulose in the pericarp, hardening the kernel and improving its
 CC handling ability. The sequences are used to produce transgenic plants
 CC used for modulating, preferably increasing, the level of the synthase
 CC in a plant cell. The plants are preferably monocots. The polynucleotide is
 CC also used as a probe or primer in the detection quantitation or
 CC isolation of gene transcripts. The probes are useful in detecting
 CC deficiencies in the level of mRNA in screenings for desired transgenic
 CC plant, for detecting mutations in the gene, for monitoring upregulation
 CC of expression or changes in enzyme activity in screening assays of the gene,
 CC compounds, for detection of any number of allelic variants of the gene,
 CC or for use as molecular markers in plant breeding programs. The
 CC isolated nucleic acids of the present invention can also be used for
 recombinant expression of their encoded polypeptides or for use as
 CC immunogens in the preparation and/or screening of antibodies. The
 CC proteins can be employed in assays for enzyme agonists or antagonists
 CC or enzyme function or for use of immunogens or antigens to obtain
 CC antibodies specifically immunoreactive with a protein.

Sequence 1074 AA;

Query Match 70.9%; Score 4024; DB 21; Length 1074;
 Best Local Similarity 69.4%; Pred. No. 0;
 Matches 749; Conservative 128; Mismatches 159; Indels 44; Gaps 10;

Db 23 DSGDPVPAPKPTKSANGQVQCQICGDTVGVSATGDFVACNECAFPCRCPEYERKEGNC 82
 Qy 62 CPOCKTRKRLKGSSPAIFGDKDEGLADEGTVEFNYPOKEK1SERMLGWHLTRKGGEEMG 121
 Db 83 CPOCKTRKRLKGSSPAIFGDKDEGLADEGTVEFNYPOKEK1SERMLGWHLTRKGGEEMG 128
 Qy 122 EPQDKEYS-----HNHLPLRSRDTSGERSAASPERLSVSTIAGKRLPYSSDV 174
 Db 129 LQGDADLSSSARHDPHHRIPRLISQGQISGEPPDASPRHSI-----RSPTSSVD 180
 Qy 175 QS--PNRIVD-----VGLGMWAKERVGDKMMKDEKTNGPVSQASERGGVQDAS 226
 Db 181 PSVFPVPVRIVDPSKDLNSYGLNSDWKERVESWVKQDKNMQLQTNKYPEARCDME--G 237
 DE 227 TDIADEALLNDBARQPLRSKVISPPSRNIPRMRVIMRLVILCFLHIRTNPVNAFA 286
 Db 238 TGSNGEDKOMVMDQARPLSPRIVLSSNINLYRILCFLHIRTNPVNAFA 287
 Qy 287 LWLWSVICBIIWFASWILUDQPKWPKPVNRETYLDRALRYDRBEGEPSOLAVIDFVSTD 346
 Db 298 LWLWSVICBIIWFASWILUDQPKWPKPVNRETYLDRALRYDRBEGEPSOLAVIDFVSTD 357
 Qy 347 PIKEPPIYTANTVLSITLAVDYPVDKVSYVEDGAMLSFESLAETSFARKWWPEKKY 406
 Db 358 PLKEPLLTANTVLSITLAVDYPVDKVSYVSDDGSMATLFESISETAEFKRWPFWCKH 417
 Qy 407 SIEPPAPEWFAAKIDYKDKVOTSFVKRRAKREBEFKRINALYSALKCPEEW 466
 Db 418 NIEPPAPEWFAAKIDYKDKVOTSFVKRRAKREBEFKRINALYSALKCPEEW 477
 Qy 467 MDQGTPWPGNNTCDDHGMNQVFLGQNGLDAEGNELPRLVY'SREKRGFOHKKAGMN 526
 Db 478 MADCTAWFCNNPRDHGMNQVFLGHSGLDGTGDNELPRLVY'SREKRGFOHKKAGMN 537
 Qy 527 ALVRSVAVLNGFILNIDCRHVNINSALREMCFLDNPNGKQCVVOFPFDGIDL 586
 Db 538 ALIVRSVAVLNGFILNIDCRHVNINSALREMCFLDNPNGKQCVVOFPFDGIDL 597
 Qy 598 NDYVANRWFEDINLRGLDGTOGPVYGTGCVFNRATLGCEPPIKVHKRPSLSKLC 646
 Db 598 HDYVANRWFEDINLRGLDGTOGPVYGTGCVFNRATLGCEPPIKVHKRPSLSKLC 646
 Qy 647 GGSRKNSKAKKEDKKKSGRHTDSTVVFNULDIEEGVEAGCPDDEKALLMSQMSLEKR 706
 Db 658 CGRRKRKNSYMSQSRSRMR-TESSAIPNFMEDIEGIE-GYEDERSVLSQRKLER 714
 Qy 707 FGGSAVVEASTLBMENGVYPPSATPENLIKAEATVIVISGYEDKSDWGMIGWIGSVEDI 766
 Db 715 FGQSPIFIATSTFMQGGIPPSNTNPASLKEATHWISCGIEDKTEWKGREGIWGVSVEDI 774
 Qy 767 LIGFKMHARGWRSYCMKLPKAFKGSAVINSLRQNLWALGSEVELSFHCPINGY 826
 Db 775 LTCRKMHARGWQSYCMKLPKAFKGSAVINSLRQNLWALGSEVELSFHCPINGY 834
 Qy 827 NGRLKFLERFAYVNTTIPITSPLLMCITLAWCLFNQFIRQISIASIWFSLELS 886
 Db 835 NGRKLLRLAYINTIVPITSPLIATCVPALCILNKTFEPISWAGNFELIFAS 894
 Qy 887 IFAMGILEMRWMSGVIDEMRNHQFWTGGVS AHLFAVFOGLIKVLAGIDTNFTVSKAS 946
 Db 895 IFATGILEMRWMSGVIDEMRNHQFWTGGVS AHLFAVFOGLIKVLAGIDTNFTVSKAS 954
 Qy 947 DEGDFEALYLFKWTLLIPPTPLILVWLVGAGVTAINGSYOSNGPFLGKFEFW 1006
 Db 955 DEGDFEALYLFKWTLLIPPTPLILVWLVGAGVTAINGSYOSNGPFLGKFEFW 1014
 Qy 1007 IVHYPFLKGLMGRQNRPPTIVVWSVLLASIFSLWWRDFTSRVTPGDIL-EGCINC 1065
 Db 1015 IHLYPFLKGLMGRONRPTIVVWSVLLASIFSLWVKIDFISPDKAAALGQCGVNC 1074

RESULT 15

ID	AYA84112	AAV84112 standard; Protein; 1074 AA.
QY	175	QS--PNRRRTVD-----VGLGVNAKERVGDGKGWKKMQEKNNGPVSTQASERGGVTDAS 226
AC	1	: : : : : : : : : : : : : : : : : : :
XX	181	PSVPVPVRTIVDPSKDLSNYSQNSDVKERVESWRVQDKNDMLQVNTRKYPEARGDME--G 237
DT	03-JUL-2000	(first entry)
DE	227	Amino acid sequence of a maize cellulose synthase.
KW	238	Maize; cellulose synthase; stalk quality; stand; silage; cellulose; transgenic plant; plant breeding marker.
OS	287	zea mays.
XX	298	WO200009706-A2.
PN	347	PKERPLPVNTVLSILANDYPVDKVSCVYFDGDAAMLSEFSLATESEFARKWVPCPKY 406
XX	358	PLKEPPPLITANTVLSILANDYPVDKVSCVYFSDDGSAMLTSEFESLSETAEFKRWWPKCKH 417
PD	407	PR 16-AUG-1999; 99WO-US18760.
PF	418	PR 17-AUG-1998; 98US-009622.
PT	467	PA (PION-) PIONEER HI-BRED INT INC.
PR	478	Pi Dhugra KS, Helentjaris TG, Bowen BA, Wang X;
XX	527	WPI: 2000-224343/19.
DR	538	N-PSDB; AAZ959506.
XX	587	New genes which encode maize cellulose synthase polypeptides in plants useful for modulating the expression of cellulose synthase in plants and to produce transgenic plants expressing the novel protein
PT	598	Claim 15; Page 126-128; 119pp; English.
PR	647	The present sequence represents a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.
XX	658	CC The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.
SQ	707	CC Sequence 1074 AA;
QY	715	Db 707 FGQSAVFAVSTLMENGVPPSATPENLKEAHLVSCGYEDKSDQMEIGWIGSYEDI 766
AC	716	Db 715 FGQSBIFIASFMTMFTOGGIPSTNPASLKEAHVSCGYEDKTENGKELGWIGSYEDI 774
CC	767	Db 767 LTGFMRHARWRSIYCMPKLFKAFKGASAPINSLDRUNQLVRALGSVEILSRHCPTWGY 826
CC	775	Db 775 LTGFMRHARWRSIYCMPKPFCKGASAPINSLDRUNQLVRALGSVEILSRHCPTWGY 834
CC	827	Db 827 NGRLKFLEPRAVYNTIYPTISIPLIMCYTLLAVLFTNFIPIOSINTSIWIWSLFLS 886
CC	835	Db 835 NGRKLKFLEPRAVYNTIYPTISIPLIMCYTLLAVLFTNFIPIOSINTSIWIWSLFLS 894
CC	887	Db 887 IFATGILEMWAWSGVGIDEWWRNEQFWVIGGSAHFAVFOGLKLVLAGIDNTFTSKAS 946
CC	895	Db 895 IFATGILEMWAWSGVGIEDEWMRNEQFWVIGGSAHFAVFOGLKLVLAGIDNTFTSKAS 954
CC	947	Db 947 DEDGDFAEILYLFKWNFTILLPTTLLIVNLVGVVAGVSYAINSGYOSWGPPLRGKLFPAFW 1006
CC	955	Db 955 DEDGDFAEILYLFKWNFTILLPTTLLIVNLVGVVAGVSYAINSGYOSWGPPLRGKLFPAFW 1014
CC	1007	Db 1007 IHLXPFLKGIMGRONRTPVWVWSVLLASIFSLWVDRPFTSRVTGDFDIL-BGINC 1065
CC	1015	Db 1015 IHLXPFLKGIMGRONRTPVWVWSVLLASIFSLWVDRPFTSRVTGDFDIL-BGINC 1074
QY	122	8PQYDEVS-----HNHLRLTSRDTSSBFSASAASPERVSVSSTAGKRLPYSSDVN 174
QY	123	83 CPOCKTTRYKROKGSVQCVQICGDTWGSATGDFVACNECAFPVCKPCYEYERKEGNQC 82
QY	62	62 CPOCKTTRYKROKGSVQCVQICGDTWGSATGDFVACNECAFPVCKPCYEYERKEGNQC 121
Db	83	83 CPOCKTTRYKROKGSVQCVQICGDTWGSATGDFVACNECAFPVCKPCYEYERKEGNQC 128
QY	122	122 8PQYDEVS-----HNHLRLTSRDTSSBFSASAASPERVSVSSTAGKRLPYSSDVN 174

Title:	AAC39336	34	103	1.8	470	1	ROCC_BAGSU	P39636 bacillus su
Perfect score:	5677	35	103	1.8	1066	1	NUC2_NEUCR	P01317 neurospora
Sequence:	1 MESEGETAGKPMKNIVPQTC. IDPFITSRVTPGPDIILECGINC 1065	36	103	1.8	1232	1	B3A3_HUMAN	P48751 homo sapien
Run on:	June 16, 2003, 11:03:13 ; Search time 15 Seconds	37	102	1.8	642	1	ARE2_YEAST	P53629 saccharomyces
	(without alignments)	38	101.5	1.8	1863	1	BRC1_HUMAN	P38398 homo sapien
	{without alignments})	39	101	1.8	617	1	VAIL_BOVIN	P31404 bos tauris
	2944.819 Million cell updates/sec	40	101	1.8	3097	1	CARD_DROME	Q15943 drosophila
		41	100.5	1.8	744	1	YTV2_CABEB	Q18600 caenorhabditis
		42	100.5	1.8	1020	1	ATC1_DROME	P22700 drosophila
		43	99.5	1.8	322	1	NUIM_STREU	P15548 strongyloides
		44	99.5	1.8	536	1	COX1_ALIMA	P80440 allomyces m
		45	99.5	1.8	714	1	HUNK_MOUSE	Q88866 mus musculus
Searched:	112892 seqs, 41476328 residues							
Total number of hits satisfying chosen parameters:	112892							
Minimum DB seq length:	0							
Maximum DB seq length:	200000000							
Post-processing:	Minimum Match 0 %							
	Maximum Match 100 %							
	Listing first 45 summaries							
Database :	SwissProt_40:*							
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.							
	SUMMARIES							
Result No.	Score	Query Match Length	DB ID	Description				
1	273	4.8	872	1 BCSA_ECO57	RESULT 1			
2	273	4.8	874	1 BCSA_SALTI	BCSA_ECO57	STANDARD;	PRT:	872 AA.
3	272	4.8	872	1 BCSA_ECOLI	08X5L7;			
4	271	4.8	874	1 BCSA_XANAC	DT	15-JUN-2002 (Rel. 41, Created)		
5	252.5	4.4	729	1 ACS2_ACEXY	DT	15-JUN-2002 (Rel. 41, Last sequence update)		
6	252.5	4.4	1596	1 ACS2_ACEXY	DT	Cellulose synthase catalytic subunit [lapp-forming] (EC 2.4.1.12).		
7	247.5	4.4	1518	1 BCA4_ACEXY	DE	BCSA OR Z4948 OR ECS4113.		
8	247.5	4.4	1518	1 BCA5_ACEXY	GN	Escherichia coli O157:H7;		
9	236.5	4.2	756	1 BCA2_ACEXY	OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
10	228.5	4.0	754	1 BCA1_ACEXY	OC	Escherichia.		
11	228	4.0	739	1 BCSA_PSEFL	RN	NCBI TaxID:83334;		
12	208.5	3.7	745	1 BCA3_ACEXY	RN	[1]		
13	196	3.5	1550	1 AGSI_ACEXY	RN	SEQUENCE FROM N.A.		
14	119	2.1	1 COX1_PARB	RC	STRAIN=O157:H7 / RIMD 050952;			
15	118.5	2.1	590	1 YKPA_CAEEL	RX	MEDLINE-21156231; PubMed-1258796;		
16	115	2.0	357	1 GOFT_HUMAN	RA	Hayashi T., Makino K., Ohnishi M., Murata T., Tanaka M., Tohe T., Iida T., Tokami H., Honda T., Sasakawa C., Kurokawa K., Ohayama K., Ogasawara N., Yasunaga T.,		
17	115	2.0	517	1 COX1_ASPIE	RA	Kuhara S., Shiba T., Hattori M., Shinagawa H.;		
18	110.5	1.9	441	1 YKDO_ECOLI	RT	"Complete genome sequence of enteropathogenic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.";		
19	110.5	1.9	672	1 A1B_HUMAN	RT	DNA RES. 8:11-22(2001).		
20	110	1.9	517	1 COX1_SRPPU	RL	P12700 paracentrot		
21	109	1.9	830	1 FARL_YEAST	CC	- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes uridine 5'-diphosphate glucose to cellulose, which is produced as an extracellular component for mechanical and chemical protection at the onset of the stationary phase, when the cells exhibit multicellular behavior (radar morphotype). Co-expression of cellulose and thin aggregative fimbriae leads to a hydrophobic network with tightly packed cells embedded in a highly inert matrix (By similarity).		
22	108	1.9	517	1 COX1_PISOC	CC	P21268 saccharomyces		
23	108	1.9	1029	1 RLP3_RAT	CC	P25001 pisaster oc		
24	107.5	1.9	1437	1 MRPS_HUMAN	CC	Ogeret6 rattus norvegicus		
25	107	1.9	3033	1 P0LG_HOVJ6	CC	O15440 homo sapien		
26	106	1.9	357	1 GSF7_CANFE	CC	P26660 genome po		
27	105.5	1.9	1076	1 YF38_SCIRPO	CC	O19133 canis familiaris		
28	105.5	1.9	4829	1 BTR6_HUMAN	CC	013892 schizosaccharomyces pombe		
29	104.5	1.8	309	1 YESP_BACSU	CC	P25001 pisaster oceanicus		
30	104	1.8	995	1 ARH7_ARARTH	CC	O9330 arabidopsis thaliana		
31	104	1.8	1167	1 CLAA_BACTU	CC	P56955 bacillus thuringiensis subsp. san miguel		
32	104	1.8	1879	1 POIN_SMV1	CC	P36286		
33	103				CC	034864 bacillus su		

ALIGNMENTS

- - PATHWAY: Bacterial cellulose biosynthesis.

- - SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential).

- - DOMAIN: There are two conserved domains in the globular part of the protein: the N-terminal domain (domain A) contains the conserved DxD motif and is possibly involved in catalysis and substrate binding. The C-terminal domain (domain B) contains the

CC substrate binding. The C-terminal domain (domain B) contains the
 CC QXRW motif and is present only in processive glycosyl
 CC transferases. It could be involved in the processivity function of
 CC the enzyme, possibly required for holding the growing glycan chain
 CC in the active site.

-1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR InterPro: IPR001173; Glycos_transf_2.
 DR Pfam: PF00535; Glycos_transf_2_1.
 KW Cellulose biosynthesis: transferase; Glycosyltransferase;
 KW Transmembrane; Inner membrane; Complete proteome.
 FT DOMAIN 271 364 CATALYTIC SUBDOMAIN A.
 FT TRANSMEM 441 501 CATALYTIC SUBDOMAIN B.
 FT TRANSMEM 30 50 POTENTIAL.
 FT TRANSMEM 151 171 POTENTIAL.
 FT TRANSMEM 173 193 POTENTIAL.
 FT TRANSMEM 230 250 POTENTIAL.
 FT TRANSMEM 525 545 POTENTIAL.
 FT TRANSMEM 547 567 POTENTIAL.
 FT TRANSMEM 592 612 POTENTIAL.
 FT TRANSMEM 634 654 POTENTIAL.
 FT TRANSMEM 668 688 POTENTIAL.
 FT TRANSMEM 833 853 POTENTIAL.
 FT ACT_SITE 313 313 POTENTIAL.
 FT ACT_SITE 457 457 POTENTIAL.
 FT SITE 360 360 SUBSTRATE BINDING (POTENTIAL).
 FT SITE 362 362 SUBSTRATE BINDING (POTENTIAL).
 SQ SEQUENCE 874 AA: MW: 33950D9CD051990 CRC64;

Query Match 4.8%; Score 273; DB 1; Length 874;
 Best Local Similarity 18.8%; Pred. No. 8.6e-12; Indels 370; Gaps 29;
 Matches 153; Conservative 91; Mismatches 201; Idents 370; Gaps 29;

OY 260 MVVIMRLVVLICLFLHYRINTNPV---PNFALWIVSVCEIWFLSWILDQEPKWPVNR 315
 OY :||:| | :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
 Db 201 MLTVISLTIVSCRYIWWRTYSTLNNDPVSLSVCGLILFAETBYAWIVLVLVIGYFOVWMLNR 260
 OY 316 ETYLDRLALRYDRGEPPSOLAALVIFVSVVDPLKEPPLYTANTVLISLAVIDYFVKVSCY 375
 OY :||:| ||:| ||:| ||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
 Db 261 QP-----VPLPKEMSQPTVDFIFVPHYN-----EDLNWVKNNTIYASGIDWPKDLNINW 309
 OY 376 VFDDGAMLSFESLAETTSFARKWVPECKYSIEPRAPEWYFAAKIDYLKDVKOTSFVWD 435
 Db 310 ILDDGGR----ESP----RQFAR---- 324
 OY 436 RRANKREYEEFKIRINALVSKALKCPEEGWVMODGTPWGNNGDHPGMQLFGLQNGSL 495
 Db 325 -----YIAR----ATHEHAKGNINNALKHA---KGFRVATFDCDHVPTSEL 370
 OY 496 DAEGNELPLRVVSYREKRCFGHQHKKAGAMNALVRSAVLTINGPILNLDCDHVINNSKA 555
 Db 330 -----YIAR----ATHEHAKGNINNALKHA---KGFRVATFDCDHVPTSEL 370
 OY 556 LREAMCFMLDPNLGKQVQVQFOPRF---DGIDKN---DRYANRNTVFFDINJLRGLGDI 608
 Db 371 QMTNGWFKE---KQLMQTPHFFSDPFERNLGRFRKTPTGELTYGLVQDGNDM 426
 OY 609 QGPYVGTGSCGVNRLHALGYVEPPRKVKRPSLSSKLCGGSRKNSKARKEDKRKSRRH 668
 Db 427 DATFFCGSCAVIRR----KP---- 442
 OY 669 TDSTVVPVNUDDIBBEGVVEGAGFDDEBKALIUMSQMSLERFGQSAVFVASTLMENGVPPSA 728
 Db 443 -----LDEI-----GGI---- 449

OY 729 TPENLILKEATHVTCISCGYEDKSDWMGMEIGWVYGSYTEDILGKUHARGHSIYONMPKLPA 788
 OY :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
 Db 450 -----AVE----- 478
 Db 479 SAGLATESTISAHIGORIRWARGMVOI-FRLDNPL---FGKGKLQAROLCYLNAMPHFLSG 534
 OY 849 IPLMVCY-----LLACLFNTQIIQI-----SNIASTWFLSLSIA 889
 OY :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
 Db 535 IPRLFITPLAELLHAYIYIAPALMIALFVPRINVHSLTNKIOGKYRHSMWEITE 594
 OY 950 GDFEABLYFKWTTLIIPPTLILIVLNVWVAGVSAINSQYQSGLPFLFLEFWIVH 1009
 OY :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
 Db 623 GGLVVEEKYVWD-VISRPYFLFLVLMNLIGVAGV-----WRYYY----- 661
 OY 1010 LYPEFLKGLMGQRQNKRPTIVVWWSULASIFSLNW 1044
 OY :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
 Db 662 -----ENELTIVY-----SLWV 675

RESULT 3
 BCSA_ECOLI STANDARD; PRT; 872 AA.
 ID BCSA_ECOLI
 AC P31653; P37654; P76712; P76713;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cellulos synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
 GN BCSA OR B3533.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=94316500; PubMed=8041620;
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes".
 RL Nucleic Acids Res. 22:2576-2586(1994).
 RL [2]
 RP REVISIONS
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=9746617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Greco J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Ma B., Shao Y.;
 RT The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP CHARACTERIZATION.
 RC STRAIN=ECOR10, ECOR12, and TOBL;
 RX Published=11260463;
 RA Zogaj X., Niintz M., Roehde M., Bokranz W., Roemling U.;
 RT "The multicellular morphotypes of *Salmonella typhimurium* and
 RT *Escherichia coli* produce cellulose as the second component of the
 RT extracellular matrix;"
 RL Mol. Microbiol. 39:1452-1463(2001).
 CC - FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
 CC uridine 5'-diphosphate glucose to cellulose, which is produced as
 CC an extracellular component for mechanical and chemical protection
 CC at the onset of the stationary phase, when the cells exhibit
 CC multicellular behavior (rدار morphology). Co-expression of
 CC cellulose and thin aggregative fibrillae leads to a hydrophobic
 CC network with tightly packed cells embedded in a highly inert

CC	-! CATALYTIC ACTIVITY: UDP-glucose + {(1,4-beta-D-glucosyl)}(N) = UDP matrix.	Db	310 ILDDGG-----	315
CC	+ {(1,4-beta-D-glucosyl)}(N+).	Qy	436 RRAMREYEFKIRINALVSKALKCPEEGWVMODGTPWPGNNTCDHPGMIOVFLGQNGL 495	
CC	-! COFACTOR: Magnesium (BY similarity).	Db	316 ---REEPROF-----	
CC	-! ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid (C-GMP).	Qy	496 DAEGELPRLVYVSREKRCFOHKKAGAMNALYRVSALTINGPFLNUCDIHYINNSKA 555	
CC	-! SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential).	Db	329 K-----YIARTT---HEHAKGNNTNALKYA---KGEPVSTFDCCAVPTRSFL 370	
CC	-! DOMAIN: There are two conserved domains in the globular part of the protein: the N-terminal domain (domain A) contains the conserved DXD motif and is possibly involved in catalysis and substrate binding. The C-terminal domain (domain B) contains the OXXRW motif and is present only in processive glycosyl transferases. It could be involved in the processivity function of the enzyme, possibly required for holding the growing glycan chain in the active site.	Qy	556 LREAMCFLMDPNLGKQCVQFFOFE--DGIDKN---DRYANRNTFFDILRGLGI 608	
CC	-! MISCELLANEOUS: The genes bscA, bcsB, bcsZ and bcsC are constitutively transcribed but cellulose synthesis occurs only when adrA, a putative transmembrane protein regulated by agfD, is expressed. Cellulose production is abolished in E.coli K12.	Db	371 QMTKGWFLKE---KQLAMQTPHFFSDPFEENLGRERKTNEGTUFGYIQQGNDW 426	
CC	-! SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.	Qy	609 QGPYVGVTGCVFNRLTYALGYEPPTRVKHKKPSSLKLCGGSRKRNKSKAKKESDKKSGRH 668	
CC	-! CAUTION: Ref.1 sequence differs from that shown due to a frameshift in position 128.	Db	427 DATFCGSCAVIRR-----KP-----	442
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	Qy	669 TDSTPVPFVNLDIIGGVEDAGFDDKEKALIMQSLEKRGQSAVFASLTIMENGGVPPSA 728	
CC	-----	Db	443 -----LDEI-----	449
CC	-----	Qy	729 TPENILKEATHVISCGYEDKSDWGMIGWIYGSVTEILDITGFKMHARGWSRSTCMKPLA 788	
CC	-----	Db	450 -----AVE-----TWEDATSLRHRGRTSAM-RIFQ 478	
CC	-----	Qy	789 FKGSPAPINISDRNLNOVLRVALGWSVILEFSRHCPIWYNGR-LKFLEFRAYVNTTYPIT 847	
CC	-----	Db	479 AAGLATELSAHIGQIRRMARGMYQI-FRDNLPL---TGKGLKFAQRLCYWNAMFHRLS 533	
CC	-----	Qy	848 SIPIMCYCTLLAVCFDTQFIPOINSNTASIWIWFSLTSIFANGILEMWWSGIDEWNR 907	
CC	-----	Db	534 GIPRILFLTAPLAFLLHAYII---YAPALMTALFV---LPHMIIASLTNSKIOGKR 585	
CC	-----	Qy	908 NEQFWVIGVSAHLAEVFGIL-----KVLAGIDTN---FWTKSKASDEDGDEAEL 957	
CC	-----	Db	586 H-SFW-----SEIYEHVLANYIAPPFLVALINPHKGKENVAK-----GGIVEBEY 630	
DR	EcoCee; EG12260; bcsA.	Qy	958 FKWTTLIPTTLVVLNVGVVAGSYAINSGIQSGWGLPFLGKLFPAFWVIVHLYFLKL 1017	
DR	InterPro: IPR01173; Glycos_transf_2.	Db	631 VDW-VISRYVFLVLUVLNVGVAGI-----WRYFYG-----	660
KW	Cellulose biosynthesis; Transferase; Glycosyltransferase;	Qy	1018 MGRONTTPITLVVWNSVLAISIFSLWV 1044	
FT	Transmembrane; Inner membrane; Complete proteome.	Db	661 -----PPT-----EMLTVVNSWW 675	
FT	DOMAIN 271 364 CATALYTIC SUBDOMAIN A.	Qy		
FT	TRANSMEM 441 501 CATALYTIC SUBDOMAIN B.	Db		
FT	TRANSMEM 150 171 POTENTIAL.	Qy		
FT	TRANSMEM 173 193 POTENTIAL.	Db		
FT	TRANSMEM 230 250 POTENTIAL.	Qy		
FT	TRANSMEM 525 545 POTENTIAL.	Db		
FT	TRANSMEM 547 567 POTENTIAL.	Qy		
FT	TRANSMEM 592 612 POTENTIAL.	Db		
FT	TRANSMEM 640 660 POTENTIAL.	Qy		
FT	TRANSMEM 668 688 POTENTIAL.	Db		
FT	TRANSMEM 833 853 POTENTIAL.	Qy		
FT	ACT_SITE 313 313 POTENTIAL.	Db		
FT	ACT_SITE 457 457 POTENTIAL.	Qy		
FT	SITE 360 360 SUBSTRATE BINDING (POTENTIAL).	Db		
SQ	SEQUENCE 362 362 SUBSTRATE BINDING (POTENTIAL).	Qy		
SQ	SEQUENCE 872 AA; 99/84 MW; 14326BBa2EB228F7 CRC64;	Db		
RESULT 4		Qy		
	BCSA_SALTY	Db		
	ID BCSA_SALTY STANDARD; PRT; 874 AA.	Qy		
Qy	931N2;	Db		
Db	15-JUN-2002 (Rel. 41, Created)	Qy		
Db	15-JUN-2002 (Rel. 41, Last sequence update)	Db		
Db	15-JUN-2002 (Rel. 41, Last annotation update)	Qy		
Db	Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).	Db		
OS	(Salmonella typhimurium)	Qy		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Salmonella.	Db		
OC	NCI_TAXID=602;	Qy		
RN	{1}	Db		
RN	SEQUENCE FROM N.A.	Qy		
RC	STRAIN=ATCC 14028;	Db		
RX	MEDLINE=21160181; PubMed=11260463;	Qy		
RA	Zogaj X., Nintz M., Rohde M., Bokranz W., Roemling U.;	Db		
RT	"the multicellular morphotypes of <i>Salmonella typhimurium</i> and <i>Escherichia coli</i> produce cellulose as the second component of the extracellular matrix."	Qy		
RT	Escherichia coli produce cellulose as the second component of the extracellular matrix."	Db		
RL	Mol. Microbiol. 39:1452-1463(2001).	Qy		
RN	[2]	Db		
RP	SEQUENCE FROM N.A.	Qy		
RC	SIRAIN-LT2;	Db		
RX	MEDLINE=21927388; PubMed=11929533;	Qy		

RA	Solano C., Garcia B., Vaille J., Berasain C., Ghigo J.-M., Gamazo C., Lasa I.;
RT	*Genetic analysis of <i>Salmonella enteritidis</i> biofilm formation: critical role for cellulose."
RT	Mol. Microbiol. 43:793-808(2002).
RL	[3]
SEQUENCE FROM N.A.	"Complete genome sequence of <i>Salmonella enterica</i> serovar Typhimurium
RT	Lt2";
RL	Nature 413:852-856(2001).
CC	-i- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes uridine 5'-diphosphate glucose to cellulose, which is produced as an extracellular component for mechanical and chemical protection at the onset of the stationary phase, when the cells exhibit multicellular behavior (rod morphotype). Co-expression of cellulose and thin aggregative fimbriae leads to a hydrophobic network with tightly packed cells embedded in a highly inert matrix.
CC	-i- CATALYTIC ACTIVITY: UDP-glucose + {(1,4-beta-D-glucosyl)(N) - UDP + [(1,4-beta-D-glucosyl)](N-1)}
CC	-i- COFACTOR: Magnesium (By similarity).
CC	-i- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid (c-di-GMP) (By similarity).
CC	-i- PATHWAY: Bacterial cellulose biosynthesis.
CC	-i- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential).
CC	-i- DOMAIN: There are two conserved domains in the globular part of the protein: the N-terminal domain (Domain A) contains the conserved DxD motif and is possibly involved in catalysis and substrate binding. The C-terminal domain (domain B) contains the QXKRW motif and is present only in processive glycosyl transferases. It could be involved in the processivity function of the enzyme, possibly required for holding the growing glycan chain in the active site.
CC	-i- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensee@isb-sib.ch).
CC	-----
DR	EMBL; AJ31570; CAC4015_1; -
DR	EMBL; AE008867; AAL2479_1; -
DR	EMBL; AJ315148; CAC80199_1; -
DR	STYGene; SG????; besa.
DR	INTERPRO; IPR01150; Cellulose_synt.
DR	INTERPRO; IPR01173; Glycos_transf_2.
PFAM; PF03552; Cellulose_synt_1.	
DR	Pfam; PF00535; Glycos_transf_2_1.
KW	Transmembrane; Inner membrane; Complete Proteome.
FT	DOMAIN 271 271 CATALYTIC SUBDOMAIN A.
FT	TRANSEM 441 501 CATALYTIC SUBDOMAIN B.
FT	TRANSEM 300 50 POTENTIAL.
FT	TRANSEM 151 171 POTENTIAL.
FT	TRANSEM 173 193 POTENTIAL.
FT	TRANSEM 230 250 POTENTIAL.
FT	TRANSEM 525 545 POTENTIAL.
FT	TRANSEM 547 567 POTENTIAL.
FT	TRANSEM 592 612 POTENTIAL.
FT	TRANSEM 634 654 POTENTIAL.
FT	TRANSEM 668 688 POTENTIAL.
FT	TRANSEM 833 853 POTENTIAL.
RA	Best Local Similarity 18.8%; Score 271; DB 1; Length 874; Matches 153; Conservative 90; Mismatches 202; Indels 370; Gaps 29;
FT	ACT_SITE 313 313 POTENTIAL.
FT	ACT_SITE 457 457 SUBSTRATE BINDING (POTENTIAL).
FT	SITE 360 360 SUBSTRATE BINDING (POTENTIAL).
FT	SITE 362 362 SUBSTRATE BINDING (POTENTIAL).
SQ	SEQUENCE 874 AA; 100043 MW: 4C9421B58606310A CRC64;
Query Match	4. 8%; Score 271; DB 1; Length 874;
Matches	153; Conservative 90; Mismatches 202; Indels 370; Gaps 29;
Qy	260 NIVMLRLVILCLFLHVRYTNV---PNAEFLWLVSVCITWFALSWILDOPKPWNVR 315
Db	201 MLIVLSLTVSCRIVWWRITSSLNWDDPPSLVCGLLFAETWAWTWLVLGFQVWPLNR 260
Qy	316 ETYLDRLALRVDREGEPSQLAVIDFVSTVPLKEPPLVTNTVLAVYQPVDVSCY 375
Db	261 QP-----VPLPKMSQWPPTDVFVPTYN--EDLVVVKNTIASLGIDWPDKLNW 309
Qy	376 VFDGAMLSEFESLAETSEFRKWPFCKKYSIEPAPWWFAAKIDLKVKVQTSFKD 435
Db	310 ILDDGR---ESF--RHFAF----- 324
Qy	436 RAMKREYEEFKRINALVSKALKCPEEGWMQDGTPWPQGINGTGDHPGMQVFLGONGGL 495
Db	325 -----HVGVH----- 329
Qy	496 DAEGNEGLPRLVYVWSREKRPGEQHKKAGAMMALVRYSAVLTNGPFTNLDCDHYINNSKA 555
Db	330 -----YIARTM---HEAKAGNTINNALKHA---KGEFAVAFDCDCHDVPTRSFL 370
Qy	556 LREACMFLMDPNLKGQVCYQVQFQRF--DGIDKN---DRYANRRTVFFDINLRGLGI 608
Db	371 QMTGMWFLKE---KOLAMMGTPHRFSPDFERNLGRFRTPNEGFLPGVQDENDMW 426
Qy	609 OGPVVYVGTCVFNRTALYGYVEPPPIKVHHKBSLLSKLGGSRKKNSKAKKSDKKSGRH 668
Db	427 DATFFGSCAVIRR-----KP----- 442
Qy	669 TDSTVPVNLDIIECVEGAGCDFDEKALLMQMSLKRFGQSAVFASTLMENGGPPSA 728
Db	443 -----LDEI-----GGI----- 449
Qy	729 TPENLIKEIAHWISCGYEDKSDWGMETGWIGWGVSVTEDLIGEKFHMARGWRSYICMPKLP 788
Db	450 -----AVB----- 478
Qy	789 FKGAPSAPINSLSPRLNQYTLRWAQGSVELLSRHCPIWYQYNGRLKFLERFAYVNTTIVITS 848
Db	479 AGLAGATESLSAHIGQRIRWARGMVOI-FRDLNDPL---FGKGKLAQRLCYINAMFHRLSG 534
Qy	849 IPLLWCT-----LLAVCLFTINQFQIPOI-----SNIASIWLSFLSIFA 889
Db	535 IPRLFLTFLATPAFLAFLLHAYIYAPALMIALFVIPHMYHASLJNSKLOGKYRHSFWSELYE 594
Qy	890 TGLEMRWWSGQGIDEWWRNEQOFWIGVGSVALHFAYQVQGLKVLLAGIDNTFTWIKASDED 949
Db	595 TVLA-----WYIARP-----LVALINPHKGKFNWK----- 622
Qy	950 GDFAEFLFKWTLIPLPTTLIVLNLGVWVAGWSVAINSGYQSWGPLFGAFAFWVIVH 1009
Db	623 GGLVKEVKYDVN-VISRPYIFVFLVNLNLGVAAVG-----WRYVGP----- 661
RESULT 5	1010 LYPFELKGLMGHQNRPTIIVVWSVILLASIFSLWV 1044
ID	BCSA_XANAC STANDARD PRT: 729 AA.
AC	P58932;
DT	15-JUN-2002 (Rel. 41, Created)
DT	15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BSA OR XAC518.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
Xanthomonas.
OX NCBI_TaxID:92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=204217;
RA da Silveira A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,
Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
Camarotte G., Canavan F., Cardoso J., Chambarro F., Ciapina L.P.,
Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
Formighieri E.F., Franco C.C., Griggio A.,
Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
Martins E.C., Medanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
Pereira H.R., Rossi A., Senna J.A.D., Silva C., de Souza F.R.,
Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.,
RT *Comparison of the genomes of two Xanthomonas pathogens with differing
RL host specificities.;
CC Nature 417:459-463(2002).
CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
uridine 5'-diphosphate glucose to cellulose, which is produced as
an extracellular component for mechanical and chemical protection
(BY similarity).
CC -1- CATALYTICITY: UDP-glucose + {(1,4-beta-D-glucosyl)}(N) = UDP
+ {(1,4-beta-D-glucosyl)}(N-1).
CC -1- COFACTOR: Magnesium (BY similarity).
CC -1- ENZYME REGULATION: Activated by bis-(3',-5') cyclic diquanylic acid
(C-di-GMP) (BY similarity).
CC -1- PATHWAY: Bacterial cellulose biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(CC (Potential)).
-1- DOMAIN: There are two conserved domains in the globular part of .
the protein: the N-terminal domain (domain A) contains
the conserved DxD motif and is possibly involved in catalysis and
substrate binding. The C-terminal domain (domain B) contains the
QXXW motif and is present only in processive glycosyl
transferases . It could be involved in the processivity function of
the enzyme, possibly required for holding the growing glycan chain
in the active site.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE01200; AAC38361.1; -
KW Cellulose biosynthesis; transferase; Glycosyltransferase;
KW Transmembrane; Inner membrane.
FT DOMAIN 151 244 CATALYTIC SUBDOMAIN A.
FT DOMAIN 321 381 CATALYTIC SUBDOMAIN B.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 405 425 POTENTIAL.
FT TRANSMEM 427 447 POTENTIAL.
FT TRANSMEM 520 540 POTENTIAL.
FT TRANSMEM 549 569 POTENTIAL.
FT TRANSMEM 610 630 POTENTIAL.
FT ACT_SITE 193 193 POTENTIAL.

FT	ACT_SITE	337	337	POTENTIAL.
FT	SITE	240	240	SUBSTRATE BINDING (POTENTIAL).
FT	SITE	242	242	SUBSTRATE BINDING (POTENTIAL).
SQ	SEQUENCE	729 AA;	80915 MW;	B9C0BB995E795B1 CRC64;
QY	Query Match	4.4%	Score 252.5;	DB 1; Length 729;
	Best Local Similarity	19.6%;	Pred. No. 2e-10;	
	Matches	153;	Conservative	85;
			Mismatches	212;
			Indels	331;
			Gaps	28;
QY	260 MVIMLRVILCFLHYRITNP--VPNA--FAWLVSVCETWFALSMILDQPKWPKPVNR	315		
Db	81 MMGMGMSLAVSCRITWRMOTMVGSAVDFTFLIGLIGAELLAFLVILVLYQVLMILN	140		
QY	316 ETYLDRLALRYDREGEPSOLAVIDFSTVDPLKEPPVANTVLISLANDYFVDKSCY	375		
Db	141 KP---VPPLADQRWPS---WDVFPIPTN--EPLSVRVTWLAISVTDMPAGKTIH	189		
QY	376 VFDDGAAMISFESIATESEFARKWVPECKYSISTERAPEWYFAKKDYLKDVKYOTSFVKD	435		
Db	190 LLDDG-----	-----		
Db	194 LQVAMGWFLRDKL---ALVQMPHYFSPDPFPERRNDTHGKVPNEGELFGQLQGNDQ	305		
QY	436 RRAMKREYEFKIRINALVSKALKCPEEGWMQMDGTWPWPGNNNTGDHPGMICQVFLGNGGL	495		
Db	195 --RRDERRAF-----CAEG-----	-----		
Db	196 DAEGNELPRLVYVSREKRPQFHKKAGAMNALVRSAVLNGPFILNLDCHYINNSKA	555		
Db	208 -----INVTRTNN-----HAKAGNINAALK---KOSGSDYVAIFDCH-IPTRSF	249		
QY	556 LREAM-CTEMLDPNKGQKVYVOPQRF-----DGIDKNDRYANRNUVFFDINLRGLD	607		
Db	250 LQVAMGWFLRDKL---ALVQMPHYFSPDPFPERRNDTHGKVPNEGELFGQLQGNDQ	305		
QY	608 IOPVYVGTGCVNTRALTYGEPPIKYKHKKPSLLSKLCGGSSRKKNKAKKEDSKKSGR	667		
Db	306 WNAATFFCCOSCAVTKTA-----	-----		
QY	668 HTDSTVPPVNLDDIEGVEGAGEDFDEKALLMSQMSLEKRGFGOSAVFVASTLNGSYPPS	727		
Db	330 -----AVE-----	-----		
Db	330 -----TVIDEATHALKLDRGRYRVAL---AVP	357		
QY	788 AFKGSPAPINLSDRNLQVRWALGSVET-----LSFRHCPTIWGYNGRLKFERAYNT	841		
QY	901 GIDEWWRERQFWWIGVSAHLRFVFOGILKVLAGIDT-NFTVSKA-----SDEDFAE	954		
Db	358 QAMGLATESLSGIVAVQRWARMQARIDNPLLGR-----GLKLSORLYLNA	407		
QY	842 TIVPTISIPLMLYCTLLAVCLFTNQFIPOISNIASIWFLSLFLSFATGI	900		
Db	408 MLHFYGVPRITYLTAPLAYLFQFGAHVIOASALMIALYALPHILQLANLTNLRVQSRFRHL	467		
QY	468 LWNEWYENTLAWTI-----FPTLVALLNPKLKGENTUPKGIVLVARSYFDQIAK	517		
QY	955 LYLFKWWTTLLIPPTTLTUVNLGVWVAG--SYAINSGYQS-WGPILGKLFCAFNVIVH	1009		
Db	518 PYLF-----LLLNVQGMWAGLRLIVGGSGEQQTW-----ENLANTLYN	559		
QY	1010 L 1010			
Db	560 M 560			
RESULT 6				
AC52_ACExY	STANDARD;	PRT;	1596 AA.	
ID AC52_ACExY				
AC 059167;				
DT 15-JUN-2002 (rel. 41, Created)				
DT 15-JUN-2002 (rel. 41, Last sequence update)				
DT 15-JUN-2002 (Rel. 41, Last annotation update)				

DE Cellulose synthase 2 [Includes: Cellulose synthase catalytic subunit
DE (UDP-forming) (EC 2.4.1.12); Cyclic di-GMP binding domain (Cellulose
SYNTHASE 2 regulatory domain)].
GN ACSAII
OS Acetobacter xylinus.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
Glucocacetobacter.
RA Glucocacetobacter.
RN NCBI_TaxID=28448;
RP [1]
SIQUENCE FROM N.A.
RC STRAINATCC 23759;
RK MSLDIN=95394846; Pubmed=7665515;
RA SAXENA I.M.; BROWN R.M. JR.;
RN Identification of a second cellulose synthase gene (acsAII) in
Acetobacter xylinum.;
RU J. Bacteriol. 177:5276-5283(1995).
CC -I- CATALYTIC ACTIVITY: UDP-glucose + [(1,4-beta-D-glucosyl)](N) = UDP
+ [(1,4-beta-D-glucosyl)](N-1).
-I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(potential).
-I- DOMAIN: There are two conserved domains in the globular part of
the catalytic subunit: the N-terminal domain (domain A) contains
the conserved DxD motif and is possibly involved in catalysis and
substrate binding. The C-terminal domain (domain B) contains the
QXRW motif and is present only in processive glycosyl
transferases. It could be involved in the processivity function of
the enzyme, possibly required for holding the growing glycan chain
in the active site.
-I-Miscellaneous: It is not essential for cellulose production in
this strain.

-I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
GLYCOSYLTRANSFERASE FAMILY 2.
-I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; U11595; AAAB05264.1; -.
DR InterPro; IPR003119; Cellosynth_A.
DR InterPro; IPR005920; Cellosynth_B.
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2_1.
DR PRINTS; PRO1439; CELUSNTHASER.
DR PRINTS; PRO1440; CELUSNTHASER.
KW Cellulose biosynthesis; transferase; Glycosyltransferase;
KW Transmembrane; Inner membrane.

FT DOMAIN 1 749 CATALYTIC.
CYCLIC DI-GMP BINDING (POTENTIAL).
CATALYTIC SUBDOMAIN A.
CATALYTIC SUBDOMAIN B.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 396 416 POTENTIAL.
FT TRANSMEM 421 441 POTENTIAL.
FT TRANSMEM 505 525 POTENTIAL.
FT TRANSMEM 544 564 POTENTIAL.
FT TRANSMEM 1553 1573 POTENTIAL.
FT ACT_SITE 331 331 POTENTIAL.
FT SITE 234 234 SUBSTRATE BINDING (POTENTIAL).
FT SITE 236 235 SUBSTRATE BINDING (POTENTIAL).
SEQUENCE 1596 AA; 175799 MW; D98A6F6259E1FJCE CRC64;

Query Match Best Local Similarity 4.4%; Score 252.5; DB 1; Length 1596;
Matches 142; Conservative 99; Mismatches 215; Indels 345; Gaps 25;

QY 261 VIMRLVLVILCFLFLWRYIT-----NPVPNAFLMLVS-----VICETWFLSWLDOF 307
Db 76 LMMMSLLSVLSLRYMWRLRLLTLECHSPQLQALSLLVVANLYALLTLCLSYFQMSWPLDKR 135
308 PKRFPVNRETYLPLRLALRDREGPSQAAVDIFVSTVPLKRPPLPVANTVLSILANDY 367
Db 136 PLPLP-----ADTTDNWVNDVYVPSN--BELSLVRSTVLGALAIDW 175
368 PVDVKSCYYFDGAGMLSFESLAETSEFARKWVFPCKKYSIEPPRPNWFAAKIDYLKDR 427
Db 175 PADKLNVYILDDG----- 188
QY 428 VQTSFVKDRRAMKRYEEFKIRINALVSKALKCBPEEGWMQDSTPPWPGNNNTGDPGMIOV 487
Db 209 ---ON----- 208
QY 548 HYNNNSKALREAMQFCFLM-DPNLGKQVCTYQFQPQR--DGIDKN---DRYANTNTFFD 599
Db 237 H-ITPRGEIJKTIGWMMAPKL---AJLQTPEHFSPPDFQNLATGQNVPBGNMYG 291
600 INLHQDLGIQGPVVGTCGVENRATLYVEPPKVKHHKPSLSSKLCGGSRKNNKSKAKKE 659
Db 292 LVQGNDNFWDATFFCGSCAIRRSRAVLGI----- 320
660 SDKIKSGRFTDSTPVFNLDIESEGVEAGFDEKDALLMSQSLERKGQSAYEVASTLM 719
Db 321 -----GGFATE----- 326
Db 327 ----- 345
Db 720 ENGGVPPSATPENILKAEATHVISGYEDKSDWGMEIGWIGYSTEDILTGKMHARGNRS 779
Db 346 AYLQQLPAA---GLSTERLMLHIGORVWRARGMLQIMRLDNPLGSG---LRNQRLCYL 399
QY 840 NTTTYPITSPILMLCTTLAVCUTCTNQFIPOLSNIASW-FLSLFLSIFATGILEMWS 898
Db 400 SAMSHFLFALPFLPRVFLASPLAFLFLGQNTIAASPFALWYAFPHVPHFHSITLSRVEGRW- 458
780 IYCPKPLPAFKGSPASPINSDRNLNOYVLRWALGSVETLFSRHCPTWYQYNGRLKFLERAVY 839
Db 899 GVGDEWWRNEQQF-VIGGSAHFAVFOGILKVLAGIDTNFWTSKASDEDDGDAELYL 957
Db 459 -----RYSFESEIYTFALFLVRYTMUMLNPKEEGENTKKGLOSSBYEFDI-- 507
958 FKWTTLILPPTLJLJVNLYGVVACV-----SYATNSGYOSWGPFLGKLFFA 1003
Db 508 ---NAVYFNLVILAVILALVRLGIGMMWYHDLALOSFALNT----- 548
QY 1004 FWVWVHLVYFLKGL-MGRNR 1023
Db 549 LWAVAVSLTIVLASIAVGRET 569

RESULT 7
BCA4_ACXY
ID BC4_ACXY
AC Q9BQJ2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative cellulose synthase 2 [Includes: Cellulose synthase
catalytic subunit [UDP-forming] (EC 2.4.1.12); Cyclic di-GMP binding
domain (Cellulose synthase 2 regulatory subunit)].
DE BCAAII-A.
GN Acetobacter xylinus.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
Glucocacetobacter.
OC NCBI_TaxID=28448;

CC	-I- PATHWAY: Bacterial cellulose biosynthesis.	Qy	495 LDABGNELPRLVYVSRERKPGFOHHKKAGMNAVLVSAVLTNGPFILNLDCHHYINNSK 554
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane	Db	207 -AENN-----HAKAGNLNLHALAV---TDSPFAVIFDCDH-VPTRG 241
CC	(Potential).	Qy	555 ALRERAMCTLM-DPNLGKCYVQFQREGDIDKNDRYANRNTYFDINLRGLDQCPVY 613
-I- DOMAIN: There are two conserved domains in the globular part of	Db	242 FLRRRTIGWMAADPN---ALQTPHRYAPDP---FORLAG--GMHVP-- 282	
the catalytic subunit; the N-terminal domain (domain A) contains	Qy	614 VGTGCVNITALYLYEPPIKVHKHPSLISKLOGGSRKNSKAKESDKKKSGRRHTSTV 673	
the conserved DXD motif and is possibly involved in catalysis and	Db	283 -PGNNMF----YGL----- 291	
substrate binding. The C-terminal domain (domain B) contains the	Qy	674 PVENLDDIBEGVREGAGFDEKALLMSOMSLEKRFQGOSAVFASLTMENGVPPSATPENL 733	
QXKRW motif and is present only in processive glycosyl	Db	292 -----VQDGND----- 311	
transfases. It could be involved in the processivity function of	Qy	734 LKEAHIVVSGYEDKSNDMGMEIIGWVIGSYVTEIDLIGEMHARWSITYCMPKLPAFKGSA 793	
the enzyme, possibly required for holding the growing glycan chain	Db	312 RREAKVMGIG-----GFATETVTDAAHTALKMORRGWGTAYLREPLA--GLA 356	
in the active site.	Qy	794 PINSDLRNQVLWALGSEVLEFSRHCPIWGYNGLKFLEPRAVNTIYIPTISIPLM 853	
CC	-I- SIMILARITY: IN THE N-TERMINAL SECTION: BELONGS TO THE	Db	357 TERLILHGQRVWARGMQLIMLDNPLMAG---JWEQRCYLSAMSHLFALIRLT 412
CC	FAMILY.	Qy	854 YCTILAVLCLFTNOFFITPOQTSNTASIWIW-SLFLSIFATGILEMRWSGVGIDEWRNEQFW 912
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	Db	413 FLYSPPLAFLGLONNIATASPLAISVIALPHIFHSVITLUSRICRW-----RSFW 462
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation	Qy	913 -VIGGGVSHLFAVEFGQTILKVLAGIDTNTVTSASDEGDFELYLFRWTTLIIPPTL 971
CC	use by non-profit institutions as long as its content is in no way	Db	463 SEIYETSLAFLFLVYRITIVLQPHKGKFNT---DKGGLLARGY-FDWA-VYPVNTLA 516
CC	modified and this statement is not removed. Usage by and for commercial	Qy	972 IWNUGVWVAGVSAINSQYQSWMPLPFGKLFPAFWWVHLYPELKGLMGRQRN 1023
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)	Db	517 GVCA2ALLRGY-FGIVMOPHDRLALQSTINTNWVSLIVIASVNGRET 568
DR	PRINTS: PRO1440: CELUSNTHASEB.	Qy	RESULT 9
KW	Cellulose biosynthesis; Transferase: Glycosyltransferase;	Db	BCA2_ACEXY
KW	Transmembrane; Inner membrane.	Qy	ID BCA2_ACEXY
FT	DOMAIN 1 731 CATALYTIC.	Db	STANDARD;
FT	DOMAIN 1518 CYCLIC DI-GMP BINDING (BY SIMILARITY).	Qy	PRT; 756 AA..
FT	DOMAIN 144 CATALYTIC SUBDOMAIN A.	Db	AC 0B2859;
FT	DOMAIN 314 CATALYTIC SUBDOMAIN B.	Qy	DT 15-JUN-2002 (Rel. 41, Created)
FT	TRANSMEM 24 POTENTIAL.	Db	DT 15-JUN-2002 (Rel. 41, Last sequence update)
FT	TRANSMEM 71 POTENTIAL.	Qy	DT 15-JUN-2002 (Rel. 41, Last annotation update)
FT	TRANSMEM 105 POTENTIAL.	Db	DE Celulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
FT	TRANSMEM 404 POTENTIAL.	Qy	GN BCSA.
FT	TRANSMEM 424 POTENTIAL.	Db	OS Acetobacter xylinus.
FT	TRANSMEM 465 POTENTIAL.	Qy	OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
FT	TRANSMEM 514 POTENTIAL.	Db	OC Gluconacetobacter.
FT	TRANSMEM 543 POTENTIAL.	Qy	OX NCBI_TaxID-28448;
FT	TRANSMEM 1481 POTENTIAL.	Db	[1]
FT	ACT_SITE 186 POTENTIAL.	Qy	RN SEQUENCE FROM N.A.
FT	ACT_SITE 330 POTENTIAL.	Db	RP STRAIN-BPR 2001;
FT	SITE 233 IMPORTANT FOR SUBSTRATE BINDING	Qy	RX MEDLINE#-98296257; PubMed#6630539;
FT	SITE 235 (POTENTIAL) IMPORTANT FOR SUBSTRATE BINDING	Db	RA Nakai T., Moriya A., Tonouchi N., Tsuchida T., Yoshinaga F.,
FT	SITE 235 (POTENTIAL) IMPORTANT FOR SUBSTRATE BINDING	Qy	RA Horimouchi S., Sone Y., Mori H., Sakai F., Hayashi T.;
SQ	SEQUENCE 1518 AA; 166464 MW; 7D7634503183DAB6 CRC64;	Db	RT "Control of expression by the cellulose synthase (bcsA) promoter
Query Match	4.4%; Score 24.7%; DB 1; Length 1518;	Qy	RT region from Acetobacter xylinum BPR 2001.";
Best Local Similarity	19.7%; Pred. No. 1.4e-09;	Db	RL Gene 213:93-100(1998).
Matches	152; Conservative 99; Mismatches 233; Indels 289; Gaps 30;	Qy	CC -I- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
Oy	261 VIMRLUVLICFLHYRITNPVNAFLWL-----VSVCIEWFALSWILQDFPKWNPVN 314	Db	CC uridine 5'-diphosphate glucose to cellulose. The thick cellulosic
Oy	315 RETYDLRALRDREGESQLAVIDTFSTVDPPLKEPLPVANTVSLAVDYPVQKWC 374	Qy	CC mats generated by this enzyme probably provide a specialized
Db	133 RR-----EHLPLPEMDAQWRSVDPVPSN--ERLSLVRKSTVLGAQDLDWPADRNV 181	CC protective environment to the bacterium (By similarity).	
Oy	375 YVFDGGMALSFSLASLSEFARKWVPFCKKYSIEPRAPEWVFAAKIDYLKDQVQSFVK 434	Db	CC -I- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl)(N) = UDP
Oy	182 YLDDG----- 187	Qy	+ ((1,4)-beta-D-glucosyl)(N).
Oy	435 DRRAMKREYEEKIRINAVALSKALKCPEGWMQDGTFPMQPGNTGDIHGMQVFLGONGG 494	Db	CC -I- COFACTOR: Magnesium (By similarity).
Oy	188 ---RRKAFHDFAVEAGA-----GYIIR----- 206	Db	CC -I- ENZYME REGULATION: Activated by bis-(3,-5') cyclic diguanlylic acid
Oy	188 ---RRKAFHDFAVEAGA-----GYIIR----- 206	Db	CC -I- PATHWAY: Bacterial cellulose biosynthesis.
Oy	188 ---RRKAFHDFAVEAGA-----GYIIR----- 206	Db	CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
Oy	188 ---RRKAFHDFAVEAGA-----GYIIR----- 206	Db	CC -I- DOMAIN: There are two conserved domains in the globular part of

CC	the catalytic subunit; the N-terminal domain (domain A) contains the conserved DXD motif and is possibly involved in catalysis and substrate binding. The C-terminal domain (domain B) contains the QXXW motif and is present only in processive glycosyltransferases. It could be involved in the processivity function of the enzyme, possibly required for holding the growing glycan chain in the active site.	QY	599 DINURGLGDIQGPVYVGCGCVFNTALGYEPPIKKHKKPSLKLGGGSRKNSRKK 658
CC	-!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.	Db	276 -----FORNLAVGYRT----- 287
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).	QY	659 ESDKKKSGRHTDSIVPVENLDDTEEGVEAGFDEKEALLMSQMSLERKRGQSAVFVASTL 718
CC	-----	Db	288 -----PEGNL--FYGVITODND-----FWDAFFCCSCA 314
CC	-----	QY	719 MENGGVPPSATPENLKEIAHVTSQGYEDKSDMGMEIOWIYGSVTEDILTGFKRMHARWR 778
CC	-----	Db	315 I-----LRREIESIG-----GFAVERTEDAHTRALMRGCGS 348
CC	-----	QY	779 SIVCWPKLUFAFKSAPINUSDRNQVNLWALSGVLEIFSRHCIIWYGYNR-LKFLEFA 837
CC	-----	Db	349 TAYL--RIVFASGLATERLITHIQMRMGIQI-FRVDNPML---GRGIGLGORLC 401
CC	-----	QY	838 YVNNTTIVPITSIPLMYCILLCLFTNFIQISNISASIPELSLFSIFNGILERW 897
DR	EMBL: AB010645; BRA3163.1; -;	Db	402 YLSAMTSFFFAIRVIFLAFLAGONITAAPLAVALAALPHMFHSIAI----- 454
DR	InterPro: IPR003919; Celisynth.A;	QY	898 SGVQIDEWARNEQFW-VIGGVSAHLFAVFOGLIKVLVLAGIDTNETVTSKA---SDEDGFA 953
DR	Pfam: PF00535; Glycos_transf_2;	Db	455 -AAKVWKGR-YSEWSEYETTMALFLYRVTIVLFLFSKGRENTEKGGVLEEEFDLG 512
PRIM'S	PR0439; CELSWHABEA;	QY	954 ELY-----LFKWHITLIPPTPLI-VNLIQVWAGVSTAINSGYQSMGPLEGKFLEFARWV 1006
KW	Cellobiose; Transferase; Glycosyltransferase; transmembrane; Inner membrane.	Db	513 ATYPNIIATIMMGGLLIGLFELIVRNFOLDVIARNAYLNCA-----WA 557
KW	-----	QY	1007 IVHYPFLKGL-MGRQR 1023
FT	TRANSMEM 167 187	Db	558 LISLILFAIAVGRETK 575
FT	TRANSMEM 409 429	-----	-----
FT	TRANSMEM 432 452	-----	-----
FT	TRANSMEM 470 490	-----	-----
FT	TRANSMEM 517 537	-----	-----
FT	TRANSMEM 517 537	-----	-----
FT	ACT_SITE 189 189	-----	-----
FT	ACT_SITE 335 335	-----	-----
FT	SITE 238 238	-----	-----
FT	SITE 240 240	-----	-----
FT	SEQUENCE 756 AA; 84561 MW; 6954F39A25E73B0A CRC64;	-----	-----
Query Match Similarity 4.2%; Score 236.5; DB 1; Length 756; Matches 166; Conservative 98; Mismatches 263; Indels 331; Gaps 34; Pred. No. 3.1e-09;	RESULT 10	-----	-----
QY	204 EKNIGPVSTQASBERGGYDADSD---ILADEALLNEAR-----OPL 244	BCAL_ACXYY	STANDARD; PRT; 754 AA.
Db	11 ESRIGRISKILSLURGASYIVGALGLCALLAATVTLNNEQLIVAACVWVFFVGRCK 70	AC	P19449; 01-FEB-1991 (Rel. 17, Created)
QY	245 SRKVSIPISSRINPYRMVITMLRVLTICLFLHYRITNPVNAFLWL-----VSVICEIW 298	DT	01-FEB-1991 (Rel. 17, Last sequence update)
Db	71 SRRYQI-----FLEVLSALVSLRVLITWLRLTLD---FNTWIOQIGVILIMAELY 119	DT	15-JUN-2002 (Rel. 41, Last annotation update)
QY	299 ALSWILDPPKWPVNRTTYLDLALRDRREGPSQLAAVDIVSTVPLKEPLVANT 358	DE	Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
Db	120 LYMLFLSYVPTIOPLHRA-----LPLPNVDDMPT---VDIPIPTY---BOLSIYRLT 168	GN	BESA.
QY	359 VLSILANDVYVDKVCYVFFDDGAMLSFESLAEATSEFARKWVPFCKYSSIEPRAPEWFA 418	OS	Acetobacter xylinum.
Db	169 VLGAIGIIMPDKVNVYIDDG----- 190	OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceas;
QY	419 AKIDYKLKDVKOTSFVKDRRAMKREYEFKRIRNALVSKALKCPEEGWMDQDTPWPWGNNT 478	OX	Gluconacetobacter.
Db	191 -----VRFPEEQ-----AKDC----- 202	NCBI_TAXID=28448;	[1]
QY	479 GDHPGMIOFWLGONGGLDPEGENELPRPLVYVRSREKRPQFGQFHKKAGAMVAVRVSAVLTG 538	RC	SEQUENCE FROM N.A., AND SEQUENCE OF 25-42 AND 196-206.
Db	203 ---GALYT-----GRVWD-----SAHAKAGNINHAIK-----RMSG 231	RX	STRAIN=1306-3; MEDLINE=91045951; Pubmed=2146681;
QY	539 PFIINLDCDHYINNSKAKLAMCFMLDPNLKGQCYVQFPQRDGDIDKNDRYANRNNTVF 598	RX	Wong H.C., Fear A.L., Calhoun R.D., Bichinger G.H., Mayer R., Amitkam D., Ben-Ziman M., Gefland D.H., Meade J.H., Emerick A.W., Bruner R., Ben-Bassat A., Tal R;
Db	232 DYLTLILDDCH--ITRAFLQIAWMWY--ADRKLALMOPTHRHSPP----- 275	RX	"Genetic organization of the cellulose synthase operon in Acetobacter xylinum"; Proc. Natl. Acad. Sci. U.S.A. 87:8130-8134(1990).
QY	479 GDHPGMIOFWLGONGGLDPEGENELPRPLVYVRSREKRPQFGQFHKKAGAMVAVRVSAVLTG 538	RN	[2] ENZYME REGULATION.
Db	191 -----VRFPEEQ-----AKDC----- 202	RP	STRAIN=1306-3;
QY	479 GDHPGMIOFWLGONGGLDPEGENELPRPLVYVRSREKRPQFGQFHKKAGAMVAVRVSAVLTG 538	RX	PubMed=1297407;
Db	191 -----VRFPEEQ-----AKDC----- 202	RA	Chang A.L., Tuckerman J.R., Gonzalez G., Mayer R., Weinhouse H., Volman G., Amikam M., Ben-Ziman M., Gilles-Gonzalez M.-A.;
QY	479 GDHPGMIOFWLGONGGLDPEGENELPRPLVYVRSREKRPQFGQFHKKAGAMVAVRVSAVLTG 538	RT	"Phosphodiesterase A1, a regulator of cellulose synthesis in Acetobacter xylinum, is a heme-based sensor.";
Db	203 ---GALYT-----GRVWD-----SAHAKAGNINHAIK-----RMSG 231	RL	-----
QY	539 PFIINLDCDHYINNSKAKLAMCFMLDPNLKGQCYVQFPQRDGDIDKNDRYANRNNTVF 598	CC	-!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes uridine 5'-diphosphate glucose to cellulose in a processive way.
Db	232 DYLTLILDDCH--ITRAFLQIAWMWY--ADRKLALMOPTHRHSPP----- 275	CC	The thick cellulosic mats generated by this enzyme probably provide a specialized protective environment to the bacterium.
QY	539 PFIINLDCDHYINNSKAKLAMCFMLDPNLKGQCYVQFPQRDGDIDKNDRYANRNNTVF 598	CC	-!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP + {(1,4)-beta-D-glucosyl}(N).
Db	232 DYLTLILDDCH--ITRAFLQIAWMWY--ADRKLALMOPTHRHSPP----- 275	CC	-!- COFACTOR: Magnesium.
QY	539 PFIINLDCDHYINNSKAKLAMCFMLDPNLKGQCYVQFPQRDGDIDKNDRYANRNNTVF 598	CC	-!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid

CC the protein; the N-terminal domain (domain A) contains
 CC the conserved DXD motif and is possibly involved in catalysis and
 CC substrate binding. The C-terminal domain (domain B) contains the
 CC OXXW motif and is present only in processive glycosyl
 CC transferases. It could be involved in the processivity function of
 CC the enzyme, possibly required for holding the growing glycan chain
 CC in the active site.

-!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.

This SWISS-Prote entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce> or send an email to license@isb-sib.ch).

DR EMBL; AV074776; AAU17842; 1;

DR Cellulose biosynthesis; Transferase; Glycosyltransferase;

KW Transmembrane; Inner membrane.

FT DOMAIN 157 250 CATALYTIC SUBDOMAIN A.

FT DOMAIN 327 387 CATALYTIC SUBDOMAIN B.

FT TRANSMEM 36 55 POTENTIAL.

FT TRANSMEM 59 76 POTENTIAL.

FT TRANSMEM 83 101 POTENTIAL.

FT TRANSMEM 116 138 POTENTIAL.

FT TRANSMEM 417 436 POTENTIAL.

FT TRANSMEM 440 462 POTENTIAL.

FT TRANSMEM 524 546 POTENTIAL.

FT TRANSMEM 551 573 POTENTIAL.

FT ACT SITE 199 199 POTENTIAL.

FT ACT SITE 343 343 POTENTIAL.

FT SITE 246 246 SUBSTRATE BINDING (POTENTIAL).

FT SITE 248 248 SUBSTRATE BINDING (POTENTIAL).

SQ SEQUENCE 739 AA: 82165 MW: 2B962EA3854B23BB CRC64;

Query Match 4.0%; Score 228; DB 1; Length 739;
 Best Local Similarity 18.8%; Pred. No. 1.3e-08; Matches 213; Indels 342; Gaps 33
 Matches 149; Conservative 87; Mismatches 213; Indels 342; Gaps 33

QY 257 PYRNVIMRLRVICG----FLHYKRTNPVNAFAWL----VSVICELWFAL---- 300

Db 80 PGRLAI-LALWVLISLVAISRYMFNLSTL--GEFTWDMFEGYGLVAAYTALIVLIFG 136

137 YVQTAWPLRRTPVNL-----KTEPEEWPTDVFIPTYN--EALSTIVK 176

QY 301 ---SWILDQOPPKPKNFVNRETYLDRLALRYDREGRSQLAADIVFYSTVDPLKEPPVTA 356

177 LTIFFAQMDWPKDPLKLRVLLDDG-----RDDDFRE---FCKRVGVN---- 215

QY 357 NTVLISLILAVDYPVVKVSCVFDDGAMLSEFSLATISEFARKWVPIKKKSIERPAEWY 416

417 FAKKIDYLKDKVQTFSVKDRRAMKREYEFEKIRINALVSAALKCPEEGWWMDGTWPGN 476

Db 216 ----- 215

QY 477 NTGDHPGMIQVFLGONGGLDAEGNELLPRLVYVSREKRPGFOHHKAGAMNALVRSAVLT 536

Db 216 -----YIRDN---NFHAKAGNLNEALKV---T 237

QY 537 NGPFILNLDCDHYINNSKALREAMCFLMDPNLGKQCYVQFPQRF---DGIDKN DRY-- 590

Db 238 DGEVTALEFDADHVPTRSFLOVSLGKFLDKKL---AMLQPDPHETSPDPFPEKFNDTERA 293

QY 591 -ANRNTVFEIDLNRGLDG1QGPVYVGTCGVFNRTALYGYPIKVKHKPKSLLSKLCGGG 649

Db 294 VPNEGELFGVLQDGNDLMMATFFGCAVIRR---EP----- 328

QY 650 RKKNSSAKKESDKKKSGRHTDSTVVFNLIDIEGVEAGAFDEDEKALLMSQMSLKERFGQ 709

Db 329 ----- 328

QY 710 SAVFVASTLMEENGVPPSATPENLILKEIAHIVISCGYEDKSQDWGMETGWIGWIGSVTEDILTG 769

Db	329	-----LLEGGV-----AVE-----	-----TWEADHTA 347
Qy	770	FRMHARGRSITCVMKPKPAFKGSAPINLSDLNQVRWALGSEVLFSSRCPWQYNGR 829	
Db	348	LKLNRLGNTAYL-AIPOAQLATSSLRSHNRQRWARGMAQTFTDNPLUGK----- 401	
Qy	830	IKFPLERAVNFTIPTISPIPLMVL-LAVCLFNQFIPQNSIASWFLSLF-SI 887	
Db	402	LKWQRICYANAMQHFFYGLPRLFLTAPLAYLFG-----AEIHPASLMIVAVLPHL 456	
Qy	888	FANGILEMRWSSVGIDWWNEOF----WIGGVSAHLFAVFOILKVLAGIDNTVT 942	
Db	457	VHSLLTSNRIQGRFRHSW-NEYETVLAWTI----LPPVVLVALVNPRAG--GFNVT 506	
Qy	943	SKASDEQDFABLYLFLKWTLLIPPTLTLNLWGVAGSYAINGQYQSGLPLEFKLF 1002	
Db	507	DKGGIDLQF----FDW-KIARPVLYLVAVNLIGLGFGIHL-----WGDASTAVTV 554	
Qy	1003	AF---WVIVL 1010	
Db	555	AINLTWFLYNL 565	
Db			
RESULT 12.			
BCA3_ACEXY			
ID		STANDARD;	PRT; 745 AA.
AC		BCA3_ACEXY	
DT	15-JUN-2002 (Rel. 41, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DE	Cellulose synthase I catalytic subunit [UDP-forming] (EC 2.4.1.12).		
GN	BCSAl.		
OS	Acetobacter xylinum.		
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;		
OC	Glucuronacetobacter;		
OK	NCBI_TAXID=28448;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-JCM 7664 / IFO 13693;		
RX	PubMed-10382968;		
RA	Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,		
RA	Inoue Y.;		
RT	"Cloning of cellulose synthase genes from <i>Acetobacter xylinum</i> JCM 7664: implication of a novel set of cellulose synthase genes.";		
RL	JCM Res. 6:109-115(1999).		
CC	DNA Res. 6:109-115(1999).		
CC	Uridine 5'-diphosphate glucose to cellulose. The thick cellulosic		
CC	matrix generated by this enzyme probably provide a specialized		
CC	protective environment to the bacterium (By similarity).		
CC	- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-beta-D-glucosyl](N) - UDP		
CC	+ [(1,4)-beta-D-glucosyl](N+1) -		
CC	- COFACTOR: Magnesium (By similarity).		
CC	- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diquanylic acid (C-di-GMP) (By similarity).		
CC	- PATHWAY: Bacterial cellulose biosynthesis.		
CC	- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential).		
CC	- DOMAIN: There are two conserved domains in the globular part of the catalytic subunit: the N-terminal domain (domain A) contains the conserved DXD motif and is possibly involved in catalysis and substrate binding. The C-terminal domain (domain B) contains the QXRW motif and is present only in processive glycosyl transferases. It could be involved in the processivity function of the enzyme, possibly required for holding the growing glycan chain in the active site.		
CC	- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial		
CC	use		

RESULT 14

COX1_PARLI	STANDARD;	PRT;	517 AA.
ID	PI2700;		
AC			
DT	01-OCT-1989 (Rel. 12, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Cytochrome c oxidase polypeptide I (EC 1.9.3.1).		
COI			
GN			
OS	Paracentrotus lividus (Common sea urchin).		
OG	Mitochondrion.		
OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoidea; Echinidae;		
OC	Paracentrotus; Echinidae;		
OX	NCBI_TAXID=7655;		
RP	[1] SEQUENCE FROM N.A.		
RX	MEDLINE=89291831; PubMed=254576;		
RA	Cantatore P., Roberti M., Rainaldi G., Gadaleta M.N., Saccone C.,		
RT	"The complete nucleotide sequence, gene organization, and genetic code of the mitochondrial genome of Paracentrotus lividus.";		
RL	J. Biol. Chem. 264:10965-10975(1989).		
RN	[2]		
RP	SEQUENCE OF 469-517 FROM N.A.		
RC	TISSUE=EGG;		
RX	MEDLINE=87248108; PubMed=3596250;		
RA	Cantatore P., Roberti M., Morisco P., Rainaldi G., Gadaleta M.N., Saccone C.,		
RT	"A novel gene order in the Paracentrotus lividus mitochondrial genome.";		
RL	Gene 53:41-54(1987).		
-1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN HEME A ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2 AND HEME A OF SUBUNIT 1 TO THE BI METALLIC CENTER FORMED BY HEME A3 AND COPPER B.			
CC	-1- CATALYTIC ACTIVITY: 4 ferrocyanochrome c + O ₂) = 4 ferricyanochrome c + 2 H ₂ O.		
CC	-1- PATHWAY: Respiratory chain; terminal step.		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL.		
CC	INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.		
CC	-1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).		
CC	-----		

RESULT 15

YKTA_CAEEL	STANDARD;	PRT;	590 AA.
ID	P34322;		
AC			
DT	01-FEB-1994 (Rel. 28, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Hypothetical protein C07A9.11 in chromosome III.		
GN	C07A9.11.		
OS	Caenorhabditis elegans.		
OC	Bukaryota; Metazoa; Nematoidea; Chromadorea; Rhabditida; Rhabditoidae; Peloderaida; Caenorhabditis.		
OX	NCBI_TAXID=6239;		
RN	[1] SEQUENCE FROM N.A.		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RX	MEDLINE=94150718; PubMed=7906398;		
RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laisteller N., Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M., Parsons J., Percy C., Riftken L., Roopra A., Saunders D., Shownkeen R., Sims M., Smalton N., Smith A., Smith M., Sonnhammer E., Staden R., Sulston J., Thillyer Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sprott J.,		

RA Wohlleben P.
 RP "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans";
 RL Nature 368:32-38(1994).
 CC !- SIMILARITY: SOME, TO YEAST YJR106W.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; 229094; CAB82341.1; -.
 DR PIR; S40707; S40707.
 DR WormPep; C07A9.11; CE00501.
 DR InterPro; IPR004837; NaCa_Exemb.
 DR Pfam; PF01099; Na_Ca_Ex; 2.
 KW Hypothetical protein.
 SQ Sequence 590 AA; 65540 MW; 71ADE3945C31BCB CRC64;
 Query Match 2.1%; Score 118.5; DB 1; Length 590;
 Best Local Similarity 19.3%; Pred. No. 0; 84; Mismatches 205; Indels 171; Gaps 26;
 Matches 111; Conservative 89; Mismatches 205; Indels 171; Gaps 26;
 QY 531 VSAVLNTNGPPFTIUNLDCDHYINNSKALRRAACMFLMDPNIGKOVQYVQFPORFDGIDKNDRY 590
 ||::|||:|||:|||:|||:|||:|||:|||:
 Db 97 VSHLILTSPPDEV-----FSTIRDLLFLVALSF-LARCFV-----F 133
 QY 591 ANRNTVFFDINRGLDIIQGPVYVGPGCVPNFKHKKPSLISKLCGGSR 650
 |||:|||:|||:|||:|||:|||:
 Db 134 YNRVILWMPMLPTFLGLYLYVITVIGAQAVHNR-----KRKA-----Q 171
 QY 651 KNS-KAKKESDKKKSSRHTDSIVPVNLDDIEEGVEGAGF-----DDEKALLMS 699
 |||:|||:|||:|||:|||:|||:|||:
 Db 172 KONSTSRSIKSLRSRSKTHSVAPMPVPETEVHDGEAPEPBISVVTGAIDLKEHMAE 231
 QY 700 QMSLEKRGQSAYVASTLMEENGVPPSATPENLKLKEAIHVISCYEDKSDMGMEICWIY 759
 |||:|||:|||:|||:|||:
 Db 232 KAQTTRRYVKRASFVNWGNGNLHYPATHNL----GISRRESELSDDEEFVVIH 285
 QY 760 GSV-----TEDILTGKMH----ARQWRSTYCMKPKAFA 789
 |:
 Db 286 GHVFOGHARESRAASLYPEPMQKSWRSRKKDTKDLAERLDPREAEDEEMNIFSKYMA 345
 QY 790 KGSAPINISDRINQVLRWALGSEELRSRHCPIWYGNGRLKELERAYVNTIYPTSI 849
 |:
 Db 346 INVVP-----NLLFKLTIPLNEMWSK-----PLT----- 370
 QY 850 PLIMCYCTULLAVCLFTNOIFI PQISNTASIWIWLSLFISIFATGILEMRWWSGVIDEWRRN 908
 |||:|||:|||:|||:
 Db 371 -LHAFTCPAFLFSIOPFLETPEFGSPGLWVYGLAVSI-VLAELLINFTELSQPKYY 428
 QY 909 EDFWWIGVS-AHLFAVFOGLKVLLAGIDTNPFTVNSKASDSDGDFEELYLFKWTTLIP 966
 |||:|||:|||:
 Db 429 ELYSYSGFIMSIAWLYLISEVNNV--TMLGVVSRVSH--VIGLTILAWS---- 476
 QY 967 PTTLLVNLVG-WVAGYSIAINSGYQSW-----GILFGKLEFAFWVWVHLYPF---L 1014
 |||:|||:|||:
 Db 477 -----NSIGDILADYS-WVKOGYPRMAMAAAGGPF-NLIMGRGL----PETIAKL 522
 QY 1015 KSLMGGRNRTPTI-VVWWSVLLASIFSLLWVRLDPF 1049
 |||:|||:|||:
 Db 523 QGKVISMINTPYRLLTFATISLATFLGIPVQKF 558

Search completed: June 16, 2003, 11:07:08
 Job time : 25 secs

				GenCore version 5.1.6
				Copyright (c) 1993 - 2003 Compugen Ltd.
OM Protein	- protein search, using sw model			
Run on:	June 16, 2003, 11:04:03 ; Search time 49 Seconds	(without alignments)		
Scoring table:	BLOSUM62	4478.371 Million cell updates/sec		
Searched:	AAC39336 5677 seqs, 206047115 residues			
Total number of hits satisfying chosen parameters:	671580			
Minimum DB seq length:	0			
Maximum DB seq length:	200000000			
Post-processing:	Minimum Match 0%			
	Maximum Match 100%			
	Listing first 45 summaries			
Database :				
	SPTRIMBLE_21; *			
1:	sp_archea:*			
2:	sp_bacteria:*			
3:	sp_fungi:*			
4:	sp_human:*			
5:	sp_invertebrate:*			
6:	sp_mammal:*			
7:	sp_mhc:*			
8:	sp_organelle:*			
9:	sp_phage:*			
10:	sp_plant:*			
11:	sp_rabbit:*			
12:	sp_virus:*			
13:	sp_vertebrate:*			
14:	sp_unclassified:*			
15:	sp_rvirus:*			
16:	sp_bacteria:*			
17:	sp_archaea:*			
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
				SUMMARIES
Result No.	Score	Query Match Length DB ID	Description	
1	5677	100.0 1065 10 048948	048948 arabidopsis	RESULT 1
2	5654	99.6 1065 10 09FHK6	09fkhk6 arabidopsis	ID 048948 PRELIMINARY; PRT: 1065 AA.
3	4954	87.3 1067 10 09XGK6	—	AC 048948;
4	4571	80.5 1077 10 Q9LJ16	Q9lj16 gossypium h	DT 01-JUN-1998 (T-EMBLrel. 06, last sequence update)
5	4564	80.4 1079 10 Q9LJ16	Q9lj16 zea mays (m	DT 01-JUN-1998 (T-EMBLrel. 06, last sequence update)
6	4529.5	79.8 1076 10 Q9LJ15	Q9lj15 zea mays (m	DT 01-JUN-2002 (T-EMBLrel. 21, last annotation update)
7	4028.5	71.0 1075 10 Q9LJ19	Q9lj19 zea mays (m	DE Cellulose synthase catalytic subunit.
8	4024	70.9 1074 10 Q9LJ16	Q9lj16 zea mays (m	GN ATH-B.
9	3984	70.8 1081 10 048946	048946 arabidopsis	OS Arabidopsis thaliana (Mouse-ear cress).
10	3900	68.7 1059 10 Q9LJ14	Q9lj14 zea mays (m	OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
11	3875.5	68.3 1086 10 Q9LJ13	Q9lj13 zea mays (m	OC Spermatophyta; Magnoliophyta; eudicots; core eudicots; Rosidae;
12	3843.5	67.7 1094 10 Q9LJ12	Q9lj12 zea mays (m	OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
13	3811	67.1 1026 10 Q9SMW6	Q9smw6 arabidopsis	OX NCBITaxID=3702;
14	3805	67.0 1026 10 Q9XH16	Q9xh16 arabidopsis	RN
15	3801	65.9 1065 10 Q9SK15	Q9sk15 arabidopsis	RP SEQUENCE FROM N.A.
16	3741	65.9 1043 10 Q9FNC3	Q9fcn3 arabidopsis	RC STRAIN=CV: COLOMBIA;
				RX MEDLINE=98111412; PubMed=9445479;
				RA Arioli T., Peng L., Betzner A.S., Burn J., Wittke W., Hertel W.,
				RA Camilleri C., Hofte H., Plazinski J., Birch R., Cork A., Glover J.,
				RA Redmond J., Williamson R.E.;
				RT "Molecular analysis of cellulose biosynthesis in Arabidopsis.";
				RL Science 279:717-720(1998).
				DR EMBL: AF027174; AAC39336; 1; -.
				DR InterPro: IPR005150; Cellulose-synt.
				DR InterPro: IPR001841; Znf-finger.
				DR Pfam: PF03552; Cellulose-synt.
				DR SMART: SM00184; RING; 1.
				SQ SEQUENCE 1065 AA; 108559 MW; 1EF01C95FB5BE4C CRC64;
				Query Match Score 5677; DB 10; Length 1065; Best Local Similarity 100.0%; Score 5677; DB 10; Length 1065; Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
				QY 1 MESEGETAGPKMKNIVPQTQICSDVNGKTVGDREVCACDICSFPVCRCPYEYERDGNO 60
				Db 1 MESEGETAGPKMKNIVPQTQICSDVNGKTVGDREVCACDICSFPVCRCPYEYERDGNO 60
				QY 61 SCPQCKTRYKRLKGSPAIPGDKDEGLADECTVFENYPQEKEISERMMLGWLTRKGEM 120

Db	61	SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM	DT	01-MAR-2001 (TREMBLrel. 16, last sequence update)
Qy	121	GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR	DE	01-JUN-2002 (TREMBLrel. 21, last annotation update)
Db	121	GERQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR	OS	Cellulose synthase catalytic subunit.
Qy	181	IVDPGVLGNVANKERVGKMKOKEKNTPVSYTQAASERGGVVIDASDILADELLNDEA	OC	Arabidopsis thaliana (Mouse-ear cress).
Db	181	IVDPGVLGNVANKERVGKMKOKEKNTPVSYTQAASERGGVVIDASDILADELLNDEA	OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Qy	241	ROPLSRKVISIPSSRINPYRMVIMRLVICLFLHYRTNPVPAFAWLVSICEIWAL	OC	Spermatophyta; Magnoliophyta; eudicots; core eudicots; Rosidae;
Db	241	ROPLSRKVISIPSSRINPYRMVIMRLVICLFLHYRTNPVPAFAWLVSICEIWAL	OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.
Qy	241	ROPLSRKVISIPSSRINPYRMVIMRLVICLFLHYRTNPVPAFAWLVSICEIWAL	OX	NEBI_TaxID=3702;
Db	361	SILAVDYPDVKSCYVFDDGAMLSFEELAESEFARKWPKCKYSIEPRAPEWYFAK	RN	{1}
Qy	301	SWILDQFPKWFPVNRETYLDRLALRYDREGEPSQLAAVDIFSTVDPLKEPLVTANTVL	RC	SEQUENCE FROM N.A.
Db	301	SWILDQFPKWFPVNRETYLDRLALRYDREGEPSQLAAVDIFSTVDPLKEPLVTANTVL	RC	STRAIN-COLINA;
Qy	421	IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD	RX	MEDLINE=99397451; PubMed=10470850;
Qy	361	SILAVDYPDVKSCYVFDDGAMLSFEELAESEFARKWPKCKYSIEPRAPEWYFAK	RA	Kaneko T., Katoh T., Saito S., Nakamura Y., Asanizu E., Kotani H.,
Db	361	SILAVDYPDVKSCYVFDDGAMLSFEELAESEFARKWPKCKYSIEPRAPEWYFAK	RA	Miyajima N., Tabata S.:
Qy	421	IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD	RT	"Structural analysis of Arabidopsis thaliana chromosome 5. IX. Sequence features of the regions of 1,011,550 bp covered by seventeen
Db	421	IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD	RT	P1 and TAC clones";
Qy	421	IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD	RL	DNA Res. 6:183-195(1999).
Db	421	IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD	DR	EMBL: AB018111; BAB09693_1; -.
Qy	481	HPGMIVQFLGONGGLDAEGNELPLRVVSYTREKPGFOHKKAGAMNALVRVSAVLNGPF	DR	DR InterPro; IPR005150; Cellulose_synt.
Db	541	ILNLDCHYINNSKALRACMFCLMDPNLGKQVYQFQREDFGKDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM	DR	DR IntPro; IPR001841; ZnfRing.
Qy	481	HPGMIVQFLGONGGLDAEGNELPLRVVSYTREKPGFOHKKAGAMNALVRVSAVLNGPF	PFam	PF03522; Cellulose_synt; 1.
Db	601	NLRLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES	SMART	SMART: SM00184; RING: 1; Cellulose_synt; 1.
Qy	601	NLRLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES	SEQUENCE	SEQUENCE 1065 AA: 119682 MW: 3AA4714CE3C4D581 CRC64;
Db	601	NLRLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES	Query	Best Local Similarity 99.8%; Score 5654; DB 10; Length 1065;
Qy	601	NLRLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES	Matches	Matches 1062; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Db	601	NLRLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES	Qy	1 MESEGEGTACKPKMNVIPQTCQICSDNVKTVDGRDFRVAACIDSFPVCPCYERKRDQNO 60
Qy	601	NLRLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES	Db	1 MESEGEGTACKPKMNVIPQTCQICSDNVKTVDGRDFRVAACIDSFPVCPCYERKRDQNO 60
Db	601	NLRLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES	Qy	121 GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180
Qy	601	NLRLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES	Db	121 GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180
Db	661	DKKSGSRITDSTVVFENLDIEGVEGAGFDDERKALLMSOMSLERKFGQSAVASTIME	Qy	61 SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120
Qy	661	DKKSGSRITDSTVVFENLDIEGVEGAGFDDERKALLMSOMSLERKFGQSAVASTIME	Db	121 GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180
Db	661	DKKSGSRITDSTVVFENLDIEGVEGAGFDDERKALLMSOMSLERKFGQSAVASTIME	Qy	61 SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120
Qy	661	DKKSGSRITDSTVVFENLDIEGVEGAGFDDERKALLMSOMSLERKFGQSAVASTIME	Db	121 GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180
Db	721	NGGVPPSATPENLLEKATHVISCGYEDKSDWGMELIGWIGWIGSVTEDIQGKMHARGRSI	Qy	61 SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120
Db	721	NGGVPPSATPENLLEKATHVISCGYEDKSDWGMELIGWIGWIGSVTEDIQGKMHARGRSI	Db	121 GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180
Db	781	YCMPLPKPAFKGSAPINSDRNLNQYRVALGWSVLFSLRCPWYGYNGRKELERFYVN	Qy	181 IVDPGVLGNVAKERVGKMKOKEKNTPVSYTQAASERGGVVIDASDILADELLNDEA 240
Qy	781	YCMPLPKPAFKGSAPINSDRNLNQYRVALGWSVLFSLRCPWYGYNGRKELERFYVN	Db	181 IVDPGVLGNVAKERVGKMKOKEKNTPVSYTQAASERGGVVIDASDILADELLNDEA 240
Db	781	YCMPLPKPAFKGSAPINSDRNLNQYRVALGWSVLFSLRCPWYGYNGRKELERFYVN	Qy	181 IVDPGVLGNVAKERVGKMKOKEKNTPVSYTQAASERGGVVIDASDILADELLNDEA 240
Qy	841	TTIPTISPIPLMCTLIACVCLFTNQITPQSNIASTWFLSPLSFATGLEMRSV	Db	241 ROPLSRKVISIPSSRINPYRMVIMRLVICLFLHYRTNPVPAFAWLVSICEIWAL
Qy	841	TTIPTISPIPLMCTLIACVCLFTNQITPQSNIASTWFLSPLSFATGLEMRSV	Db	241 ROPLSRKVISIPSSRINPYRMVIMRLVICLFLHYRTNPVPAFAWLVSICEIWAL
Db	841	TTIPTISPIPLMCTLIACVCLFTNQITPQSNIASTWFLSPLSFATGLEMRSV	Qy	361 SILAVDYPDVKSCYVFDDGAMLSFEELAESEFARKWPKCKYSIEPRAPEWYFAK 420
Qy	901	GIDEMWRNRQEWVWVGGYSAHLFAVFOGTLKLVLAGIDNFNTVSKASDEDGFAYELFW	Db	361 SILAVDYPDVKSCYVFDDGAMLSFEELAESEFARKWPKCKYSIEPRAPEWYFAK 420
Db	901	GIDEMWRNRQEWVWVGGYSAHLFAVFOGTLKLVLAGIDNFNTVSKASDEDGFAYELFW	Qy	421 IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD 480
Qy	961	TTLLIPPTPLTINLVGAGYSYATNSGQSWGPLGKLFCAFVWVHLVYPLKGLMR	Db	421 IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD 480
Db	961	TTLLIPPTPLTINLVGAGYSYATNSGQSWGPLGKLFCAFVWVHLVYPLKGLMR	Qy	481 HPGMIVQFLGONGGLDAEGNELPLRVVSYTREKPGFOHKKAGAMNALVRVSAVLNGPF 540
Qy	1021	QRRTPTIWVWSVLASLFLSLWWRDFTSRVTPGDLLEGINC 1065	Db	481 HPGMIVQFLGONGGLDAEGNELPLRVVSYTREKPGFOHKKAGAMNALVRVSAVLNGPF 540
Db	1021	QRRTPTIWVWSVLASLFLSLWWRDFTSRVTPGDLLEGINC 1065	Qy	541 ILNLDCHYINNSKALRACMFCLMDPNLGKQVYQFQREDFRVAACIDSFPVCPCYERKRDQNO 600
RESULT	2		Db	541 ILNLDCHYINNSKALRACMFCLMDPNLGKQVYQFQREDFRVAACIDSFPVCPCYERKRDQNO 600
O9FHK6	ID	PRELIMINARY;	Db	601 NLRGLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES 660
O9FHK6	AC	PRT; 1065 AA.	Db	601 NLRGLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES 660
DT	01-MAR-2001 (TREMBLrel. 16, last sequence update)			
DE	01-JUN-2002 (TREMBLrel. 21, last annotation update)			
OS	Cellulose synthase catalytic subunit.			
OC	Arabidopsis thaliana (Mouse-ear cress).			
RC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
RC	Spermatophyta; Magnoliophyta; eu dicots; core eu dicots; Rosidae;			
RC	eu rosids II; Brassicales; Brassicaceae; Arabidopsis.			
RC	eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
RC	Kaneko T., Katoh T., Saito S., Nakamura Y., Asanizu E., Kotani H.,			
RC	Miyajima N., Tabata S.:			
RC	"Structural analysis of Arabidopsis thaliana chromosome 5. IX. Sequence features of the regions of 1,011,550 bp covered by seventeen			
RT	Pl and TAC clones";			
RT	DNA Res. 6:183-195(1999).			
RL	EMBL: AB018111; BAB09693_1; -.			
DR	DR InterPro; IPR005150; Cellulose_synt.			
DR	DR IntPro; IPR001841; ZnfRing.			
DR	DR PF03522; Cellulose_synt; 1.			
DR	DR SMART; SM00184; RING; 1; Cellulose_synt; 1.			
DR	DR SEQUENCE 1065 AA: 119682 MW: 3AA4714CE3C4D581 CRC64;			
Query	99.8%	Score 5654; DB 10; Length 1065;		
Best Local Similarity	99.8%	Score 5654; DB 10; Length 1065;		
Matches	1062	Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
Qy	1	MESEGEGTACKPKMNVIPQTCQICSDNVKTVDGRDFRVAACIDSFPVCPCYERKRDQNO 60		
Db	1	MESEGEGTACKPKMNVIPQTCQICSDNVKTVDGRDFRVAACIDSFPVCPCYERKRDQNO 60		
Qy	61	SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120		
Db	61	SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120		
Qy	61	SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120		
Db	61	SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120		
Qy	121	GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180		
Db	121	GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180		
Qy	121	GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180		
Db	121	GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180		
Qy	181	IVDPGVLGNVAKERVGKMKOKEKNTPVSYTQAASERGGVVIDASDILADELLNDEA 240		
Db	181	IVDPGVLGNVAKERVGKMKOKEKNTPVSYTQAASERGGVVIDASDILADELLNDEA 240		
Qy	241	ROPLSRKVISIPSSRINPYRMVIMRLVICLFLHYRTNPVPAFAWLVSICEIWAL		
Db	241	ROPLSRKVISIPSSRINPYRMVIMRLVICLFLHYRTNPVPAFAWLVSICEIWAL		
Qy	301	SWILDQFPKWFPVNRETYLDRLALRYDREGEPSQLAAVDIFSTVDPLKEPLVTANTVL		
Db	301	SWILDQFPKWFPVNRETYLDRLALRYDREGEPSQLAAVDIFSTVDPLKEPLVTANTVL		
Qy	361	SILAVDYPDVKSCYVFDDGAMLSFEELAESEFARKWPKCKYSIEPRAPEWYFAK		
Db	361	SILAVDYPDVKSCYVFDDGAMLSFEELAESEFARKWPKCKYSIEPRAPEWYFAK		
Qy	421	IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD 480		
Db	421	IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD 480		
Qy	481	HPGMIVQFLGONGGLDAEGNELPLRVVSYTREKPGFOHKKAGAMNALVRVSAVLNGPF 540		
Db	481	HPGMIVQFLGONGGLDAEGNELPLRVVSYTREKPGFOHKKAGAMNALVRVSAVLNGPF 540		
Qy	541	ILNLDCHYINNSKALRACMFCLMDPNLGKQVYQFQREDFRVAACIDSFPVCPCYERKRDQNO 600		
Db	541	ILNLDCHYINNSKALRACMFCLMDPNLGKQVYQFQREDFRVAACIDSFPVCPCYERKRDQNO 600		
Qy	601	NLRGLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES 660		
Db	601	NLRGLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES 660		
Qy	601	NLRGLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES 660		
Db	601	NLRGLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES 660		
DT	01-MAR-2001 (TREMBLrel. 16, last sequence update)			
DE	01-JUN-2002 (TREMBLrel. 21, last annotation update)			
OS	Cellulose synthase catalytic subunit.			
OC	Arabidopsis thaliana (Mouse-ear cress).			
RC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
RC	Spermatophyta; Magnoliophyta; eu dicots; core eu dicots; Rosidae;			
RC	eu rosids II; Brassicales; Brassicaceae; Arabidopsis.			
RC	Kaneko T., Katoh T., Saito S., Nakamura Y., Asanizu E., Kotani H.,			
RC	Miyajima N., Tabata S.:			
RC	"Structural analysis of Arabidopsis thaliana chromosome 5. IX. Sequence features of the regions of 1,011,550 bp covered by seventeen			
RT	Pl and TAC clones";			
RT	DNA Res. 6:183-195(1999).			
RL	EMBL: AB018111; BAB09693_1; -.			
DR	DR InterPro; IPR005150; Cellulose_synt.			
DR	DR IntPro; IPR001841; ZnfRing.			
DR	DR PF03522; Cellulose_synt; 1.			
DR	DR SMART; SM00184; RING; 1; Cellulose_synt; 1.			
DR	DR SEQUENCE 1065 AA: 119682 MW: 3AA4714CE3C4D581 CRC64;			
Query	99.8%	Score 5654; DB 10; Length 1065;		
Best Local Similarity	99.8%	Score 5654; DB 10; Length 1065;		
Matches	1062	Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
Qy	1	MESEGEGTACKPKMNVIPQTCQICSDNVKTVDGRDFRVAACIDSFPVCPCYERKRDQNO 60		
Db	1	MESEGEGTACKPKMNVIPQTCQICSDNVKTVDGRDFRVAACIDSFPVCPCYERKRDQNO 60		
Qy	61	SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120		
Db	61	SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120		
Qy	61	SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120		
Db	61	SCPOCKTRYKRKGSPAFIGDKDGEDGLADEGTEFNPQEKEKISERMGWHLTRGKEM 120		
Qy	121	GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180		
Db	121	GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180		
Qy	121	GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180		
Db	121	GEPQDYKEVSHNHLPLRTSRODTSGEFSASASPERLSYSSTIAGKRKLYPSSDNQSPNRR 180		
Qy	181	IVDPGVLGNVAKERVGKMKOKEKNTPVSYTQAASERGGVVIDASDILADELLNDEA 240		
Db	181	IVDPGVLGNVAKERVGKMKOKEKNTPVSYTQAASERGGVVIDASDILADELLNDEA 240		
Qy	241	ROPLSRKVISIPSSRINPYRMVIMRLVICLFLHYRTNPVPAFAWLVSICEIWAL		
Db	241	ROPLSRKVISIPSSRINPYRMVIMRLVICLFLHYRTNPVPAFAWLVSICEIWAL		
Qy	301	SWILDQFPKWFPVNRETYLDRLALRYDREGEPSQLAAVDIFSTVDPLKEPLVTANTVL		
Db	301	SWILDQFPKWFPVNRETYLDRLALRYDREGEPSQLAAVDIFSTVDPLKEPLVTANTVL		
Qy	361	SILAVDYPDVKSCYVFDDGAMLSFEELAESEFARKWPKCKYSIEPRAPEWYFAK		
Db	361	SILAVDYPDVKSCYVFDDGAMLSFEELAESEFARKWPKCKYSIEPRAPEWYFAK		
Qy	421	IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD 480		
Db	421	IDYLKDQVTSFKDRRAMKREVEEKFIRINALVSKALKCPERGWMODGTWPGNNITGD 480		
Qy	481	HPGMIVQFLGONGGLDAEGNELPLRVVSYTREKPGFOHKKAGAMNALVRVSAVLNGPF 540		
Db	481	HPGMIVQFLGONGGLDAEGNELPLRVVSYTREKPGFOHKKAGAMNALVRVSAVLNGPF 540		
Qy	541	ILNLDCHYINNSKALRACMFCLMDPNLGKQVYQFQREDFRVAACIDSFPVCPCYERKRDQNO 600		
Db	541	ILNLDCHYINNSKALRACMFCLMDPNLGKQVYQFQREDFRVAACIDSFPVCPCYERKRDQNO 600		
Qy	601	NLRGLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES 660		
Db	601	NLRGLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES 660		
Qy	601	NLRGLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES 660		
Db	601	NLRGLDGQGPVWVGTGCVFNRTALYGYEPPIKVHKKPSLISKLCGGSRSRKNKAKKES 660		

Oy	661	DRKSGRGTDSYPVFNUDDIEBGEVAGFDDEKALLSQMSLEKRGOSAVVASTIME	720	QY	61	SCPOCKTRYKULGSPAIPIGKEDGLADEGVF---NYPKEKISERMLGWHLJRG	116
Db	661	DKRSGRITDSTPVFNIDIEBGEVAGFDDEKALLSQMSLEKRGOSAVVASTIME	720	Db	61	SCPCKTRYKKQKGSPAILGDRETGGDADCGASDFYSENOEQOKLAERQGWNAKYR	120
Qy	721	NGGVPSATPENILKEATHVISCGYEDKSDWGMIGWLYGSVTEIDLTCFKMARGWNSI	780	Qy	117	GEEMGPQYDKEVSHHLPRTSRDTSGERSAASPERLSVST-TAGKKLUPSYSSDVQ	175
Db	721	NGGVPSATPENILKEATHVISCGYEDKSDWGMIGWLYGSVTEIDLTCFKMARGWNSI	780	Db	121	GEDVGAPTYDKEISHNHIPLLTSQEVSGELSAASPERLSHASPVGAGK-----	170
Qy	781	YCPKLPFKGSPAINISDRNLNQYLRWALGSVTLFSRHCPIWGYNRLKEFRAVN	840	Qy	176	SPNRLTDPV---GLGNWAKERVGDWKMKQERNTPYST-QAASERGCVIDASTD	228
Db	781	YCPKLPFKGSPAINISDRNLNQYLRWALGSVTLFSRHCPIWGYNRLKEFRAVN	840	Db	171	-SSRVDPVPEFGSSGLGNWAKERVGDWKMQEKTVPSTCQATSERGLDIASTD	229
Qy	841	TIVYPTSIPLMCTULLAVCLETNQFIPOISNIASWFLSFLSLTATGLEMWSGV	900	Qy	229	LDAEALNDARQPLSKYVSIPSSRINPYRVMIRLVLICFLAHYRITNPVNAFLW	288
Db	841	TIVYPTSIPLMCTULLAVCLETNQFIPOISNIASWFLSFLSLTATGLEMWSGV	900	Db	230	VLDDSLNDARQPLSKYVSVSSSRINPYRVMIRLVLICFLAHYRITNPVNAFLW	289
Qy	901	GIDEWWRNQFWIGVSAHLFAVFGQGILKVLAGIDNTNTVTSKASDEDGDFABLFKRW	960	Qy	289	DVSVCETIWELSLWIDQFQPKWPVRETYDLALRYDRGEGPSLAUDVIFSVTDP	348
Db	901	GIDEWWRNQFWIGVSAHLFAVFGQGILKVLAGIDNTNTVTSKASDEDGDFABLFKRW	960	Db	350	KEPPLVANTVLSILAVDYPDKVSIVYSDGAMTEFEASSETSEFARKVYFCKNYI	409
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	409	EPRAPEWYFAAKIDYLKDVKYOTSFVDRRANKREYEFKINALYVSKALCPEEGHMQ	468
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	410	EPRAPEWYFAOKIDYLKDVKYOTSFVDRRANKREYEFKVLINGLYAKAQVPEEGHMQ	469
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	469	DGTPWPGNNCDHPGMQVFGQNGLDAESENPLPVYVREKRCFGFOHKKAGANAL	528
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	470	DGTPWPGNNCDHPGMQVFGQNGLDAESENPLPVYVREKRCFGFOHKKAGANAL	529
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	529	VRVASVLTNGBFTLNIJDCHVHINNSKALREACFLMDPNLQRKQVCFOPRFDGIDND	588
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	530	VRVASVLTNGAFLNEDCHVHINNSKALREACFLMDPNLQRKQVCFOPRFDGIDND	589
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	589	RYANRIVTFEINRLEDGTOGPVYVGTGCVFNRTALYGYRPPIKYKKKSKULLSKCUGG	648
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	590	RYANRIVTFEIDNLRGLDGIQGPVYVGTGCVFNRTALYGYRPPIKYKKKSKULLSKCUGG	649
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	649	SRRKNSK-AKESDKKSGRITDSTPVFNIDIEBGEVAGFDDEKALLSQMSLEKRF	707
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	650	SRKKSSKSKGSKDKKSKGKHVDSTPVFNIDIEBGEVAGFDDEKALLSQMSLEKRF	709
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	708	QOSAVVASTIMENGVPPSATPENILKEATHVISCGYEDKSDWGMIGWLYGSVTEIDL	767
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	710	QOSAVVASTIMENGVPPSATPENILKEATHVISCGYEDKSDWGMIGWLYGSVTEIDL	769
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	768	TGFKMARGWISIYCKPLPKPFKGSPAINISDRNLNQYLRWALGSVTLFSRHCPIWGY	827
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	770	SRKKSSKSKGSKDKKSKGKHVDSTPVFNIDIEBGEVAGFDDEKALLSQMSLEKRF	709
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	828	GRKELERFAYNTTIPSIPLMCTULLAVCLETNQFIPOISNIASWFLSFLSLTATGLEMWSL	887
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	830	GRKLWERFAYNTTIPSIPLMCTULLAVCLETNQFIPOISNIASWFLSFLSLTATGLEMWSL	889
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	888	FATGILKMKWNGVQDOWMAREQFWIGVSAHLFAVFGQGILKVLAGIDNTVTKASD	947
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	890	FATGILKMKWNGVQDOWMAREQFWIGVSAHLFAVFGQGILKVLAGIDNTVTKASD	949
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	948	EDGDFAEYLKWTWILLIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWV	1007
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	950	EDGDFAEYLKWTWILLIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWV	1009
Qy	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Qy	1008	WLYPPLKGMLGRQNPTPTVWSILLASFLSLWRIDFTSRVGDPLECGINC	1065
Db	961	TIVIIPPTIPLTIVLNGLVYVAGVSYAINGYQSGPFLERKLEFATWVTHLYPELGKMR	1020	Db	1010	IHYPLFLKGMLGRQNPTPTVWSILLASFLSLWRIDFTSRVGDPLECGINC	1067

Qy	4 EGEGTAG--KPMKNIVPOTCQICSDNNGKTVDGDRFVACDICSPCPYCIEVERDKGNQS	61	Db	83 CPOCKTTRYKRKGSPRUGHGDEEDEV-DDLNEFENYKQGS-----GKPEWO 128
Qy	1 : : : : : : : : : : : : : : : : : : :	62 EPOYDKEVS-----HNHLRLTSRDTSEFSASPERIUSVSTIAGGRKLPIVSSDVN 174	Qy	122 EPOYDKEVS-----HNHLRLTSRDTSEFSASPERIUSVSTIAGGRKLPIVSSDVN 174
Db	654 FSSLGGG-RKKTSKSKSSEKKSHRHADSVFVNLDIEGEGSQFDDELSLMSQM 712	Db	129 LQGDDADLSSSARHEPHRHRFLRTSGQQISEIPDASPRHSI-----RSPTSYVD 180	
Qy	702 SLEKRGFGOSAVFASTLMENGGVPPSATPENLKEIAHVISCGYEDKSDWGMETGWIGS 761	Qy	175 QS---ENRRIVD-----VGJGNVAKERYGKWMQEKNGKPVSTOAASERGGYDIDAS 226	
Db	713 SLEKRGFGOSAVFASTLMENGGVPPSATPESLKEIAHVISCGYEDKTDWGTGTEIWIGS 772	Db	181 PSVPVRIVDPSKDLNSYGLNSVGDNSVWKRKDNMQVTNKYPEARG-DME-G 238	
Qy	762 VHEDLITGPKMHARGWSIYCMPLPAFKGSAPINSLRNLQVLRWALGSVELFSRCIP 821	Qy	227 TDILADEALNDEAOPLSRKVSISPSRINPRYIMRLVLYLCLEHYRHTNPVNAFA 286	
Db	773 VHEDLITGPKMHARGWSIYCMPLPAFKGSAPINSLRNLQVLRWALGSIELTSRCP 832	Db	239 TGSNGEXDMQMDARLPLSRVPI:TSQNOLNLYRVVILRLTILCFFOYRSHPVDAYG 298	
Qy	822 IWGYNGRKLKFLERFAYNTTIPITSIPLMCTLLAVCLFTNQFIPQISNIASTWEL 881	Qy	287 IWLVLSVTCIETWFALSMWLDOPTPKWFVNRETYLDRLALRYREGEPSQLAAVIDFSTVD 346	
Db	833 IWGYGGRLKFLERFAYNTTIPITSIPLMCTLLAVCLFTNQFIPQISNIASTWEL 892	Db	299 IWLVLSVTCIETWFALSMWLDOPTPKWFVNRETYLDRLALRYREGEPSQLAAVIDFSTVD 358	
Qy	882 SIFLISIATGILEMRWSQVGIDERMWRNRNQFWVIGGSVAHLFAVFOGLKVLVLAGTTNTV 941	Qy	347 PLKEPPLYVNTLSTLAVDOPDKVSCYVFDDGAMLSPESTAESEFAKWKWPCKY 406	
Db	893 SLEISIATGILEMRWSQVGIDERMWRNRNQFWVIGGSVAHLFAVFOGLKVLVLAGTTNTV 952	Db	359 PLKEPPLYVNTLSTLAVDOPDKVSCYVFDDGAMLSPESTAESEFAKWKWPCKY 418	
Qy	942 TSKASDGDGFABLYLKWTWTLIPPTPLLIVLNLYVGWAGVSAINSYQSMOPLGKUF 1001	Qy	407 SIPERAPWEEPAKIDYLKDVKWQTSVKDRRAMKREYEEFRKRALVSKALKCPBEGW 466	
Db	953 TEKATDEGDFDABLYLKWTWTLIPPTPLLIVLNLYVGWAGVSAINSYQSMOPLGKUF 1012	Db	419 NIEPRAPEFEFAQIDYLKQKLIQPSVKEERRAMKREYEEFRKRALVSKALKCPBEGW 478	
Qy	1002 FAWFWVHLYPLFKLGMQKRNPTIVVWMAILSITFSLMWVRIIDPFTTRVGPDIAC 1061	Qy	467 MDQGTPWPGNNNTGDPGMIOVFLGONGGLDREGNELPRLVSVSREKPGRGRHKKAGMN 526	
Db	1013 FAWFWVHLYPLFKLGMQKRNPTIVVWMAILSITFSLMWVRIIDPFTTRVGPDIAC 1072	Db	479 MADGTAWPGNNPRDHGPMLQVFLGHSGGLDIDGNELPLRVYVSREKPGRGRHKKAGMN 538	
Qy	1062 GINC 1065	Qy	527 ALVRYSAVLTNGPTELNLDDHYINNSKAIREAMPLMDPNLGKQCVYOPFOPREDGDK 586	
Db	1073 GINC 1076	Db	539 ALIRYSAVLTNGPTELNLDDHYINNSKAIREAMCMMDPALGRKTCYOPFOPREDGDK 598	
RESULT 7				
Qy	QYL19 PRELIMINARY; PRT; 1075 AA.	Db	587 NDYANRNTVFEINRGLGQIOPGVYVGCFVNRTALYGEPEPKVKHKPSLISK 646	
ID	QYL19;	Db	599 HDYANRNTVFEINRGLGQIOPGVYVGCFVNRTALYGEPEPKVKHKPSLISK 658	
AC	01-OCT-2000 (TREMBLrel. 15, Created)	Qy	647 GGSRKNSKAKKESDKKKSGHHTDSTVPUFVNLDDEBVGAGFDEDEKALLMSQLEKR 706	
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)	Db	659 CGRRKKKKNSIMDSQRSLRMR-TESSAPIFMNEDIEEGIE--GYEDERSYLNMSORKLEKR 715	
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	Qy	707 FGOSAVFASTLMENGGVPPSATPENLKEIAHVISCGYEDKSDWGMETGWIGS 766	
DE	Cellulose synthase-1.	Db	716 FGQSPLEFIASTPMTGQPIPSTNPASLKEIAHVISCGYEDKTEWKEIGHVYGSYTDI 775	
GN	CESA-1.	Qy	767 LTGFKMHARGWSIYCMPLPAFKGSAPINSLRNLQVLRWALGSVELFSHCPIWGY 826	
OS	zea mays (Maize).	Db	776 LTGFKMHARGWSIYCMPLPAFKGSAPINSLRNLQVLRWALGSVELFSHCPIWGY 835	
OC	Eukaryota; Virchiplantae; Streptophyta; Embryophyta; Tracheophyta;	Qy	827 NGRLKFLERFAYNTTIPITSIPLMCTLLAVCLFTNQFIPQISNIASTWLSLFLS 886	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCI clade;	Db	836 NGRLKLFLERFAYNTTIPITSIPLMCTLLAVCLFTNQFIPQISNIASTWLSLFLS 895	
OX	Panicoidae; Andropogoneae; Zea.	Qy	887 IFATGILEMRWSQVGIDERMWRNRNQFWVIGGSVAHLFAVFOGLKVLVLAGTTNTV 946	
NCBI-TAXID=4577;	[1]	Db	896 IFATGILEMRWSQVGIDERMWRNRNQFWVIGGSVAHLFAVFOGLKVLVLAGTTNTV 955	
RN	SEQUENCE FROM N_A.	Qy	947 DEDGDEAELYLKWTWTLIPPTPLLIVLNLYVGWAGVSAINSYQSMOPLGKUF 1006	
RQ	MDLINE=20:98328; PubMed=10938350;	Db	956 DEDGDEAELYLKWTWTLIPPTPLLIVLNLYVGWAGVSAINSYQSMOPLGKUF 1015	
RX	Holland N., Holland D., Relentless T., Dhuga K.S.,	Qy	1007 IVHLYPLFKLGMGRONTVPTIVVWMAILSITFSLMWVRIIDPFTSRVGPDIAC 1065	
RA	Xocostle-Cazares B., Delmer D.P.;	Db	1016 ILHLYPLFKLGMGRONTVPTIVVWMAILSITFSLMWVRIIDPFTSRVGPDIAC 1075	
RA	A comparative analysis of the plant cellulose synthase (Cesa) gene family.",	Qy		
RT	Plant Physiol. 123:1313-1324 (2000).	Db		
RT	EMBL: AF200525; RAAF8961-1; -.	Qy		
RT	InterPro: IPR001510; Cellulose_synt.	Db		
DR	Inter-Pro: IPR001841; Znf_ring.	Qy		
DR	Pfam: PF03552; Cellulose_synt; 1.	Db		
DR	SMART: SM00184; RING; 1.	Qy		
SQ	SEQUENCE FROM N_A.	Db		
Qy	Query Match 71.0%; Score 4028.5; DB 10; Length 1075; Best Local Similarity 69.6%; Pred. No. 4e-304; 13; Indels 45; Gaps 12; Matches 752; Conservative 130; Mismatches 12; SEQUENCE 1075 AA; 121181 MW; 67A4FB07A811F33 CRC64;	Db		

QY	995	PLRGLLFEEFWVIVHLYPPLKGMLGRQNRPPTIVVVWSVLASIFSLWWRIDPFTSRV	1054	QY
Db	1015	PLAGKLEFAFWVIVHLYPELKGLVGRQRNRPPTIVVVWSVLASIFSLWWRIDPFTLAKD	1074	QY
QY	1055	GPDTECGNC 1065		QY
Db	1075	GPLIECGLDC 1085		QY
RESULT 12				
Q9LL12		PRELIMINARY;	PRT;	1094 AA.
AC	Q9LL12;			
DT	01-OCT-2000	(TREMBLrel. 15, Created)		
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)		
DE	Cellobiose synthase-8.			
GN	CSSA-8			
OS	zea mays (Maize)			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;			
OC	Panicoidae; Andropogoneae; zea.			
OX	NCBI_TAXID=4577;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20398328; PubMed=10938300;			
RA	Holland N., Holland D., Helentjaris T., Dhuga K.S.,			
RA	Koconoostle-Cazzares B., Delmer D.P.;			
RT	"A comparative analysis of the plant cellulose synthase (Cesa) gene family";			
RL	Plant Physiol. 123:1313-1324 (2000).			
DR	EMBL: AF20532; AAF89368.1; -.			
DR	InterPro: IPR05150; Cellulose_synt.			
DR	InterPro: IPR0184; Znf_firing.			
DR	Pram: PF03522; Cellulose_synt; 1.			
DR	SMART: SM00184; RING: 1.			
SQ	SEQUENCE FROM N.A.			
Query Match	67.7%	Score 3843.5; DB 10; Length 1094;		
Best Local Similarity	66.3%	Pred. No. 1e-289;		
Matches	727; Conservative	134; Mismatches	178; Indels	57; Gaps
QY	2 ESEGETAGPKPMKIVPQTQICSDNGKTVGDRFVACDCISFPVCPOYEYERKDGNOS	61		
QY	62 CPOCKTRYKRLKGSPAPQDKDGLAD-----EGTVEFNTPQEKEISERMGWHL	113		
QY	83 CPOCKTRYKRLKGSPAPQDKDGLAD-----EGTVEFNTPQEKEISERMGWHL	113		
Db	114 RGKGEEMGEPQDYKEYSHNLPRLTSQL--DTSGEFSASASPERVSNSSTIAGKR--L	167		
QY	141 YGRGD-AHPGFSPV--PNVPLITNGQVADDEEDGVDDELEGEFGQDGAHEDPQ--VVAESMRQMS	140		
Db	168 PYSSDNQSPNRRVDP-----VGLGNWAWKERDVGWIKMOKERNNTGPVSTQAASERGGV	221		
QY	195 PF-ADPNLPVQPSMDPSKDLAAYGGYSTAWKERMEGWKOER---LQHVRSEGG	247		
Db	222 DIPASTDILADEALDEAROPLSRKSVPSRSPRNPYRATVIMLUVILFLHYTRINPV	281		
QY	248 DWGDD--ADLPIM-DEARQPUSRKVPISSRNPYRATVIMLUVILFLHYTRINPV	303		
Db	282 PNFAFLWVSYCEIWFAWSLWDOFPKWFVNRETYLRLAFLYDREGEPSQLAANDIF	341		
QY	304 KDAFLWLJISVICEWIFWAMSWILDOFPKWLPIERTYLRLSRLRDKEGQPSLAPIDFF	363		
QY	342 VSTVDPKLEPPVANTVVISLAVDYPWDVKVSCWVFDDGAAMLSEFSLAETSEFARKWP	401		
Db	402 FCKYSISRRAPENYFAAKIDYLUKQYDFVKVORRAMEREYEFKIRNALYSAKLCP	461		
QY	424 FSKRFNIEPRAPEWYFQKIDYLUKQVAAFSVFRERAMREYEFKIRNALVAKQCP	483		
Db	364 VSTVDPKLEPPVANTVVISLAVDYPWDVKVSCWVFDDGAAMLSEFSLAETSEFARKWP	423		
RESULT 13				
Q9SWW6		PRELIMINARY;	PRT;	1026 AA.
ID	Q9SWW6			
AC	Q9SWW6;			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)		
DE	Cellulose synthase catalytic subunit.			
DE	IRX3 OR TB6B_80.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid III; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CV_LANDSBURG_ERECTA;			
RC	MEDLINE=92564300; PubMed=10330464;			
RA	Taylor N.G., Scheible W.R., Cutler S., Somerville C.R., Turner S.R.;			
RT	"The irregular xylem3 locus of Arabidopsis encodes a cellulose synthase required for secondary cell wall synthesis.;"			
RT	Plant Cell 11:769-780(1999).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RA	Sato S., Nakamura Y., Kaneko T., Kato T., Asanizu E., Kotani H.,			
RA	Tabata S., Mewes H.W., Rudd S., Lemcke K., Meyer K.F.X.;			
RA	Submitted (AUG 2000) to the EMBL/GenBank/DDBJ databases.			

RN [3]	SEQUENCE FROM N. A.	Db 791 KLUKWLERFAYANTITYPFTSIPPLIAYCILPAICLCLTDKFIMPPISTFASLFFISLENSII 850
RR EU Arabidopsis sequencing project;	DR 889 ARGILEMRMSGVGDIEWRNEQEWIGGVSALHFAVFOGLIKVLLAGDTNFNTWSKASDE 948	
RA Submitted (AUG-2000) to the EMBL/GenBank/DDJB databases.	DR EMBL; AF091713; ADA0885; 1; -;	
RL DR EMBL; AL39142; CAC01737; 1; -;	DR InterPro; IPR005150; Cellulose_synt.	
DR IntePro; IPR00184; Znf_fing.	DR Pfam; PF03552; Cellulose_synt; 1.	
DR SMART; SM00184; RING; 1; -;	DR SEQUENCE 1026 AA; 115797 MW; 503BFBC78BE6E511 CRC64;	
SQ Matches 721; Conservative 115; Mismatches 143; Indels 98; Gaps 17;	Query Match 67.1%; Score 3811; DB 10; Length 1026; Best Local Similarity 66.9%; Pred. No. 3..1e-207; Matches 721; Conservative 115; Mismatches 143; Indels 98; Gaps 17;	
Oy 10 KPMKNIVPQTQTCSDNVGKTYDGRFVACDTCSPVCPYCEYERKGNDGSCPQCTRY 69	Db 27 KKLNUQGQFCICGDOIGLTVEGDLFACNCACGFPACRPCYEYERKGNDGSCPQCTRY 86	
Oy 70 KLRKGSPPAIPGKDDEGLAD--EGTVFENYPOKEKISERMIGWHLTRGKGEMGE-PQY 125	Db 87 KLRKGSPRVEGDEDIDDDIEFEENIEHQDKHHSAAEAMLYGKMSYGRGPEDDENGFR 146	
Oy 126 DKEVSHNLPRUTSRQDSGEASAASPERLSVSSTAGG-----KRL-PISSDVNQ 175	Db 147 PRVTAGH-----SGEF-----PVGGYNGEHHGHLHKRVPIPPSEAG 184	
Oy 176 SPNRRTVPVGGLGNVANKERVGDCKMKQEKNTGPVSTQASERGVDIDASTILADEAL 235	Db 185 SGG-----GWERMDWKL-QHGNIGPEP-----DDDPENGLI-----216	
Oy 236 LNDEAROPLSKRKVISPRSSRINPYRMVIMRLYTLFLHYRITNPVNAFLWLSERVICE 295	Db 217 -DEAROPLSKRKVIASSKINPYRMVIAWRVLVILAVFLYRLNPHDALGJLWLSERVICE 274	
Oy 296 TWFLSWLQDPKPWFVNREYTLRDLALRDREGESQLAADIYTFVSTVOPKEPLVT 355	Db 275 IEFAVSWLQDPKPWFPIERETYLDRSLRVEGEPMNLAVIDPVFVSTVOPKEPLVT 334	
Oy 356 ANTVLSILANDVYDVKSYVFDGGAMLSFLSBLAESTSEFARKWVIFCKKKSIEPRAPEW 415	Db 335 SWVLSLJAMDYPVEKSCCVYSDGASMFLTEESLSEAFARKWVWPCKKSIEPRAPEM 394	
Oy 416 YFAAKTDLKDKVOTSEVKDRRMRKREYEEFKIRINALVSKALKCPEBGMWQDGTPWPG 475	Db 395 YFLKVKDLQDKVHPFVKERRAKMREYEEFKVRINAQVAKASKVPLEGWIMQDGTPWPG 454	
Oy 476 NNTGDDHGMIOFLGONGGLDAGEGNEPLRVVYSEKKRGFOHKKAGAMNALVYSAVL 535	Db 455 NWTKDHGMIOFLGHSGGFDVEGHELPFRLVVSREKRGFOHKKAGAMNALVYAVGL 514	
Oy 536 TNQPFILNDCHYINNSKALREAMCILMPALGKQVYQFQREDFGDKNDRYANRNT 595	Db 515 TNAPFMLNDCHYINNSKALREAMCILMPALGKQVYQFQREDFGDKNDRYANRNT 574	
Oy 596 VFFDINIRGLDTQGPVYVGTCVENVNTRALCYEPPTRVKHKRKPSSLKLCG----GSR 650	Db 575 VFFDINKKLOGIQGPVYVGTCVFRQALGYEPP--KGPKRKPMS--GCCPCFGRR 630	
Oy 651 KNSKAKKRESKKRSGRHTDSTPVUNLDDIEBGGG--AGFDEKALLMSOMSLERFGQ 709	Db 631 RKKKK-----FSKNDMGDVALLGAGDKHELMSEMFKEFGQ 670	
Oy 710 SAVFVASTLMEENGVPPSATPENLKEAHIWVSGKEDKSDMGMETGWWIYGSVTEILTG 769	Db 770 FKMHARGRSVSYCMPLKPAFKGSAPNLISDRNQVIRWALGSVLFSSRHCPIWGY-NG 828	
Oy 770 FKMHARGRSVSYCMPLKPAFKGSAPNLISDRNQVIRWALGSVLFSSRHCPIWGY-NG 828	Db 731 FRMHCRGWSIYCMPLKPAFKGSAPNLISDRNQVIRWALGSVLFSSRHCPIWGYGG 790	
Oy 829 RLKFLERFAYVNTIPTSIPLMCLTLLAVCLFTNQFIQOISNIAISIWSLFLSLIF 888	Db 217 -DEAROPLSKRKVIASSKINPYRMVIAWRVLVILAVFLYRLNPHDALGJLWLSERVICE 274	
Db 296 TWFLSWLQDPKPWFVNREYTLRDLALRDREGESQLAADIYTFVSTVOPKEPLVT 355	Db 275 IEFAVSWLQDPKPWFPIERETYLDRSLRVEGEPMNLAVIDPVFVSTVOPKEPLVT 334	
Db 909 HLYPFLKSLMGRONRTPTIIVWWSVLLASIFSLWVDRPFTSRVCGPDLICGINC 1065	Db 970 HLYPFLKSLMGRONRTPTIIVWWSVLLASIFSLWVDRPFTSRVCGPDLICGINC 1026	
RESULT 14	Db 909XHP6 PRELIMINARY; PRT; 1026 AA.	
Oy 10 KPMKNIVPQTQTCSDNVGKTYDGRFVACDTCSPVCPYCEYERKGNDGSCPQCTRY 69	Db 909XHP6 PRELIMINARY; PRT; 1026 AA.	
Oy 236 LNDEAROPLSKRKVISPRSSRINPYRMVIMRLYTLFLHYRITNPVNAFLWLSERVICE 295	Db 909XHP6 PRELIMINARY; PRT; 1026 AA.	
Oy 296 TWFLSWLQDPKPWFVNREYTLRDLALRDREGESQLAADIYTFVSTVOPKEPLVT 355	Db 909XHP6 PRELIMINARY; PRT; 1026 AA.	
Oy 356 ANTVLSILANDVYDVKSYVFDGGAMLSFLSBLAESTSEFARKWVIFCKKKSIEPRAPEW 415	Db 335 SWVLSLJAMDYPVEKSCCVYSDGASMFLTEESLSEAFARKWVWPCKKSIEPRAPEM 394	
Oy 416 YFAAKTDLKDKVOTSEVKDRRMRKREYEEFKIRINALVSKALKCPEBGMWQDGTPWPG 475	Db 395 YFLKVKDLQDKVHPFVKERRAKMREYEEFKVRINAQVAKASKVPLEGWIMQDGTPWPG 454	
Oy 476 NNTGDDHGMIOFLGONGGLDAGEGNEPLRVVYSEKKRGFOHKKAGAMNALVYSAVL 535	Db 455 NWTKDHGMIOFLGHSGGFDVEGHELPFRLVVSREKRGFOHKKAGAMNALVYAVGL 514	
Oy 536 TNQPFILNDCHYINNSKALREAMCILMPALGKQVYQFQREDFGDKNDRYANRNT 595	Db 515 TNAPFMLNDCHYINNSKALREAMCILMPALGKQVYQFQREDFGDKNDRYANRNT 574	
Oy 596 VFFDINIRGLDTQGPVYVGTCVENVNTRALCYEPPTRVKHKRKPSSLKLCG----GSR 650	Db 575 VFFDINKKLOGIQGPVYVGTCVFRQALGYEPP--KGPKRKPMS--GCCPCFGRR 630	
Oy 651 KNSKAKKRESKKRSGRHTDSTPVUNLDDIEBGGG--AGFDEKALLMSOMSLERFGQ 709	Db 631 RKKKK-----FSKNDMGDVALLGAGDKHELMSEMFKEFGQ 670	
Oy 710 SAVFVASTLMEENGVPPSATPENLKEAHIWVSGKEDKSDMGMETGWWIYGSVTEILTG 769	Db 770 FKMHARGRSVSYCMPLKPAFKGSAPNLISDRNQVIRWALGSVLFSSRHCPIWGY-NG 828	
Oy 770 FKMHARGRSVSYCMPLKPAFKGSAPNLISDRNQVIRWALGSVLFSSRHCPIWGY-NG 828	Db 731 FRMHCRGWSIYCMPLKPAFKGSAPNLISDRNQVIRWALGSVLFSSRHCPIWGYGG 790	
Oy 829 RLKFLERFAYVNTIPTSIPLMCLTLLAVCLFTNQFIQOISNIAISIWSLFLSLIF 888	Db 217 -DEAROPLSKRKVIASSKINPYRMVIAWRVLVILAVFLYRLNPHDALGJLWLSERVICE 274	
Db 296 TWFLSWLQDPKPWFVNREYTLRDLALRDREGESQLAADIYTFVSTVOPKEPLVT 355	Db 275 IEFAVSWLQDPKPWFPIERETYLDRSLRVEGEPMNLAVIDPVFVSTVOPKEPLVT 334	

QY 758 IYGSYTEDILTGFKMHARGRSIVCMPKPAFKGSAPINLSDRLNQVRLAALGSVEILS 817
Db 760 IYGSYTEDILTGFKMHARGNSIVCVPSPAFKGSAPINLSDRLNQVRLAALGSIEILS 819
QY 818 RHCPIWYGNGRKLTERAYVNNTIPTPSIPLIMCYTULAVCLFTNQETIPQNSIAS 877
Db 820 RHCPIWYGNGRKLTERAYVNNTIPTPSIPLIMCYTULAVCLFTNQETIPQNSIAS 879
Qy 878 IWPLSLFLSIFATGLEMMSGVGDEWMARNEQFWIGVSAHLEAVFOGLIKVLAGIDT 937
Db 880 LCFMLFASIVASALEKVASDVALEDWARNEQFWIGVSAHLEAVFOGLIKVLAGIDT 939
Qy 938 NFTVSKASDEDDGDFAEYLQFWTTLLIPPTTLIVNLGVWAGVSYATNGYOSWGLPF 997
Db 940 NFTVSKASDEDDGDFAEYLQFWTSLLIPPTTLIVNLGVWAGVSYATNGYOSWGLPF 999
Qy 998 GKLFFAFWVTVHLYPFLKIGMGRONTPTVWMSVLLASIFSLWVRDPFTSEVTG 1055
Db 1000 GKLFFAFWVVAHLYPFLKIGLGRONTPTVWMSVLLASIFSLWVRINPFS-TTG 1056

Search completed: June 16, 2003, 11:07:59
Job time : 55 secs

THIS PAGE BLANK (USPTO)

(USPTO)